

高分子の内部構造を探れ!

最適化モデルを用いた実験データの解析

田中未来

Group Epsilon Common Person

2015 年 6 月 14 日
Group Epsilon 2015 3rd Meeting

発表の流れ

- 1 準備: 凸 2 次最適化問題とその解法
- 2 導入: フォトクロミック分子と量子収率分布
- 3 測定データと反応速度式
- 4 量子収率分布推定のための最適化モデル
- 5 数値実験

発表の流れ

- 1 準備: 凸 2 次最適化問題とその解法
- 2 導入: フォトクロミック分子と量子収率分布
- 3 測定データと反応速度式
- 4 量子収率分布推定のための最適化モデル
- 5 数値実験

凸 2 次最適化問題と最適性条件

2 次最適化問題

$$\left| \begin{array}{l} \text{minimize} \quad \frac{1}{2} \mathbf{x}^\top \mathbf{Q} \mathbf{x} + \mathbf{c}^\top \mathbf{x} \\ \text{subject to} \quad \mathbf{A} \mathbf{x} = \mathbf{b} \\ \mathbf{x} \geq \mathbf{0} \end{array} \right.$$

\mathbf{Q} : 半正定値 \Rightarrow 目的関数は凸 \Rightarrow 主双対内点法により多項式時間可解

$$\forall \mathbf{x}, \mathbf{x}^\top \mathbf{Q} \mathbf{x} \geq 0$$

注: \mathbf{Q} : 半正定値でない \Rightarrow NP 困難

\mathbf{x} : 最適解 \iff 次の不等式系を満たす \mathbf{y}, \mathbf{s} が存在 (双対定理)

$$\begin{aligned} \mathbf{A} \mathbf{x} &= \mathbf{b} & \mathbf{x} &\geq \mathbf{0} \\ \mathbf{A}^\top \mathbf{y} + \mathbf{s} &= \mathbf{Q} \mathbf{x} + \mathbf{c} & \mathbf{s} &\geq \mathbf{0} \\ \mathbf{x} \circ \mathbf{s} &= \mathbf{0} \end{aligned}$$

最適化問題を解く \Rightarrow 不等式系を解く

最適性条件と中心パス

最適性条件

$$\begin{aligned} \mathbf{Ax} &= \mathbf{b} & \mathbf{x} &\geq \mathbf{0} \\ \mathbf{A}^\top \mathbf{y} + \mathbf{s} &= \mathbf{Qx} + \mathbf{c} & \mathbf{s} &\geq \mathbf{0} \\ \mathbf{x} \circ \mathbf{s} &= \mathbf{0} \end{aligned}$$

中心パス: 最適性条件をパラメータ $\mu > 0$ で擾動

$$\begin{aligned} \mathbf{Ax} &= \mathbf{b} & \mathbf{x} &\geq \mathbf{0} \\ \mathbf{A}^\top \mathbf{y} + \mathbf{s} &= \mathbf{Qx} + \mathbf{c} & \mathbf{s} &\geq \mathbf{0} \\ \mathbf{x} \circ \mathbf{s} &= \mu \mathbf{1} \end{aligned}$$

性質:

- $\forall \mu > 0, \exists! (\mathbf{x}_\mu, \mathbf{y}_\mu, \mathbf{s}_\mu)$ s.t. $\mathcal{C}_\mu(\mathbf{x}_\mu, \mathbf{y}_\mu, \mathbf{s}_\mu)$ holds
- $(\mathbf{x}_\mu, \mathbf{y}_\mu, \mathbf{s}_\mu) \rightarrow (\mathbf{x}^*, \mathbf{y}^*, \mathbf{s}^*)$ as $\mu \downarrow 0$

主双対内点法

中心パス

$$\begin{aligned} \mathbf{Ax} &= \mathbf{b} & \mathbf{x} &\geq \mathbf{0} \\ \mathbf{A}^\top \mathbf{y} + \mathbf{s} &= \mathbf{Qx} + \mathbf{c} & \mathbf{s} &\geq \mathbf{0} \\ \mathbf{x} \circ \mathbf{s} &= \mu \mathbf{1} \end{aligned}$$

探索方向: 中心パスの式を 1 次近似して得られる方向

$$\begin{aligned} \mathbf{A}(\mathbf{x} + \Delta \mathbf{x}) &= \mathbf{b} \\ \mathbf{A}^\top (\mathbf{y} + \Delta \mathbf{y}) + (\mathbf{s} + \Delta \mathbf{s}) &= \mathbf{Q}(\mathbf{x} + \Delta \mathbf{x}) + \mathbf{c} \\ \mathbf{x} \circ \mathbf{s} + \Delta \mathbf{x} \circ \mathbf{s} + \mathbf{x} \circ \Delta \mathbf{s} + \cancel{\Delta \mathbf{x} \circ \Delta \mathbf{s}} &= \mu \mathbf{1} \end{aligned}$$

ステップサイズ: 非負制約を満たす

$$\alpha^* = \max\{\alpha > 0 : \mathbf{x} + \alpha \Delta \mathbf{x} \geq \mathbf{0}, \mathbf{s} + \alpha \Delta \mathbf{s} \geq \mathbf{0}\}$$

反復点の更新: 探索方向にステップサイズだけ進む

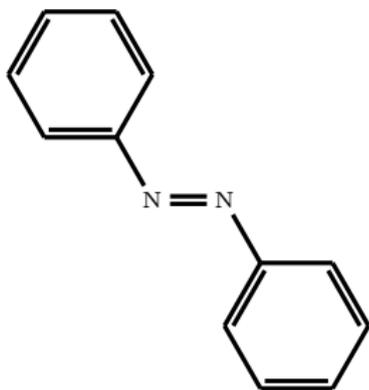
$$(\mathbf{x}, \mathbf{y}, \mathbf{s}) := (\mathbf{x}, \mathbf{y}, \mathbf{s}) + (1 - \epsilon) \alpha^* (\Delta \mathbf{x}, \Delta \mathbf{y}, \Delta \mathbf{s})$$

発表の流れ

- 1 準備: 凸 2 次最適化問題とその解法
- 2 導入: フォトクロミック分子と量子収率分布
- 3 測定データと反応速度式
- 4 量子収率分布推定のための最適化モデル
- 5 数値実験

フォトクロミック分子

- 光化学反応により変色する分子
- 光ディスクなどの記憶材料に応用可能



trans-アゾベンゼン (無色)



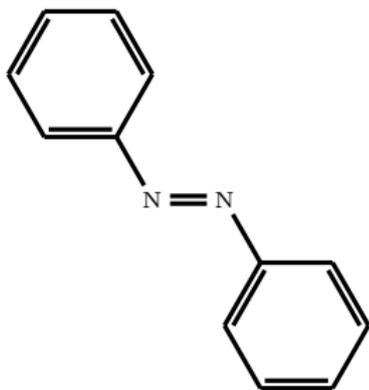
cis-アゾベンゼン (橙色)

量子収率

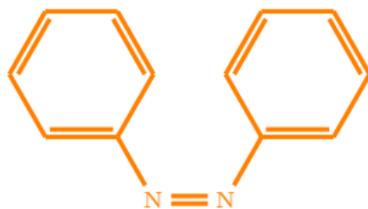
定義: ある光化学反応における量子収率 ϕ

$$\phi = \frac{\text{反応した分子の数}}{\text{吸収した光子の数}}$$

なんとなくの意味: “光化学反応の起こりやすさ”



trans-アゾベンゼン (無色)



cis-アゾベンゼン (橙色)

量子収率分布

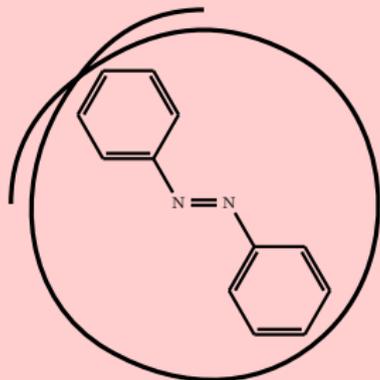
フォトクロミック分子 in 溶液

溶液中ではフォトクロミック分子は自由に動ける

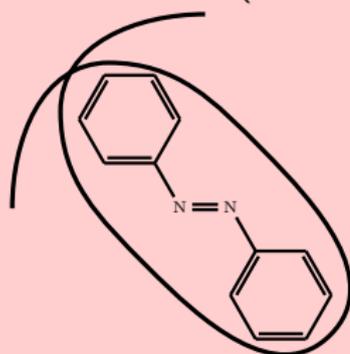
⇨ 量子収率は分子固有の一定値をとる

フォトクロミック分子 in 高分子固体

量子収率は高分子固体内の自由体積に依存した分布をもつ (量子収率分布)



量子収率: 大



量子収率: 小

高分子固体の量子収率 or 自由体積の測定法いろいろ

量子収率 \propto 自由体積 \approx 高分子鎖の間隙の大きさ

フォトクロミックプローブ法

- 高分子固体中にフォトクロミック分子を分散
- 光化学反応の速度から内部の量子収率を推定

ガス吸着法

- 高分子固体にガスを吹き付ける
- ガスの吸着量から表面の自由体積を推定

陽電子消滅法

- 陽電子を高分子固体に入射する
- 消滅するまでの時間から内部の自由体積を推定

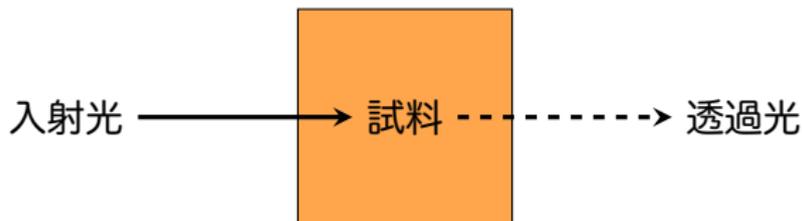
発表の流れ

- 1 準備: 凸 2 次最適化問題とその解法
- 2 導入: フォトクロミック分子と量子収率分布
- 3 測定データと反応速度式**
- 4 量子収率分布推定のための最適化モデル
- 5 数値実験

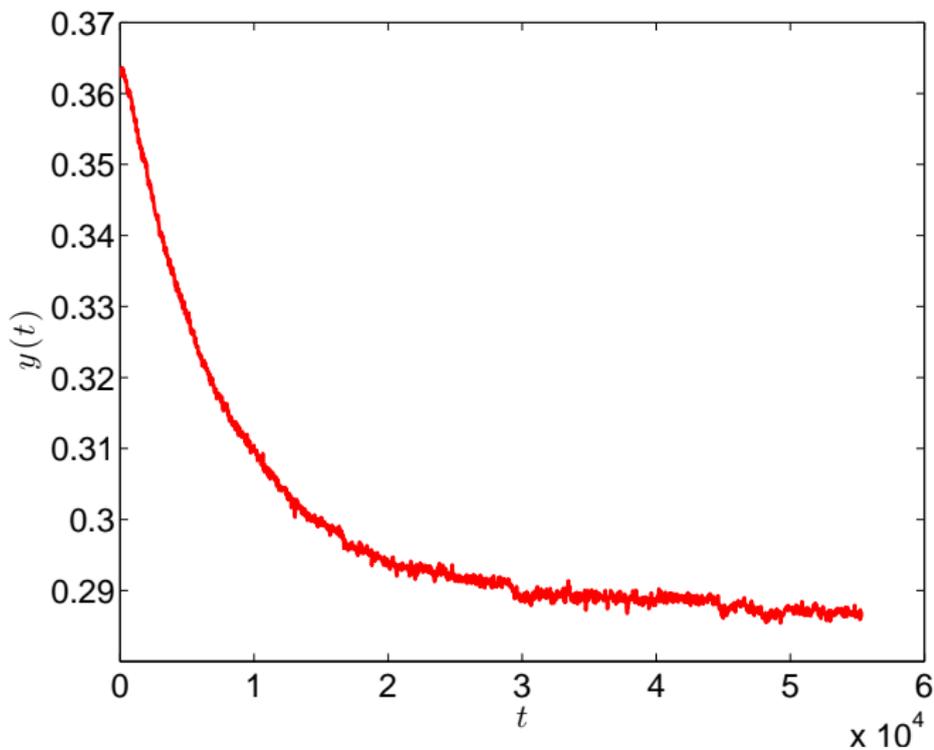
実験: 吸光度の時間変化の測定

フォトクロミック分子を分散させた高分子固体に光を照射

⇒ 吸光度 = $\log_{10} \left(\frac{\text{入射光の強度}}{\text{透過光の強度}} \right)$ の時間変化を測定



- 高分子固体: ポリメタクリル酸メチル
- フォトクロミック分子: アゾベンゼン
- 照射光波長: 365 nm (紫外光)

吸光度の測定データ $y(t)$ 

2 成分系の反応速度式

量子収率が一定のときの trans 体の減少速度式

$$\begin{aligned}
 -\frac{d}{dt}C_{\text{trans}}(t)Sl = & \\
 & I_0S(1 - 10^{-X(t)})\frac{\epsilon_{\text{trans}}C_{\text{trans}}(t)l}{\epsilon_{\text{trans}}C_{\text{trans}}(t)l + \epsilon_{\text{cis}}C_{\text{cis}}(t)l}\phi_{\text{trans}\rightarrow\text{cis}} \\
 & - I_0S(1 - 10^{-X(t)})\frac{\epsilon_{\text{cis}}C_{\text{cis}}(t)l}{\epsilon_{\text{trans}}C_{\text{trans}}(t)l + \epsilon_{\text{cis}}C_{\text{cis}}(t)l}\phi_{\text{cis}\rightarrow\text{trans}}
 \end{aligned}$$

時刻 t における吸光度 $X(t)$

$$X(t) = \epsilon_{\text{trans}}C_{\text{trans}}(t)l + \epsilon_{\text{cis}}C_{\text{cis}}(t)l$$

ただし

- I_0 : 入射光線強度, l : 光路長, S : サンプル表面積
- $C_{\text{trans}}, C_{\text{cis}}$: trans 体, cis 体のモル濃度
- $\epsilon_{\text{trans}}, \epsilon_{\text{cis}}$: trans 体, cis 体のモル吸光係数
- $\phi_{\text{trans}\rightarrow\text{cis}}, \phi_{\text{cis}\rightarrow\text{trans}}$: 各反応の量子収率

吸光度に関する微分方程式

気合で整理すると次を得る:

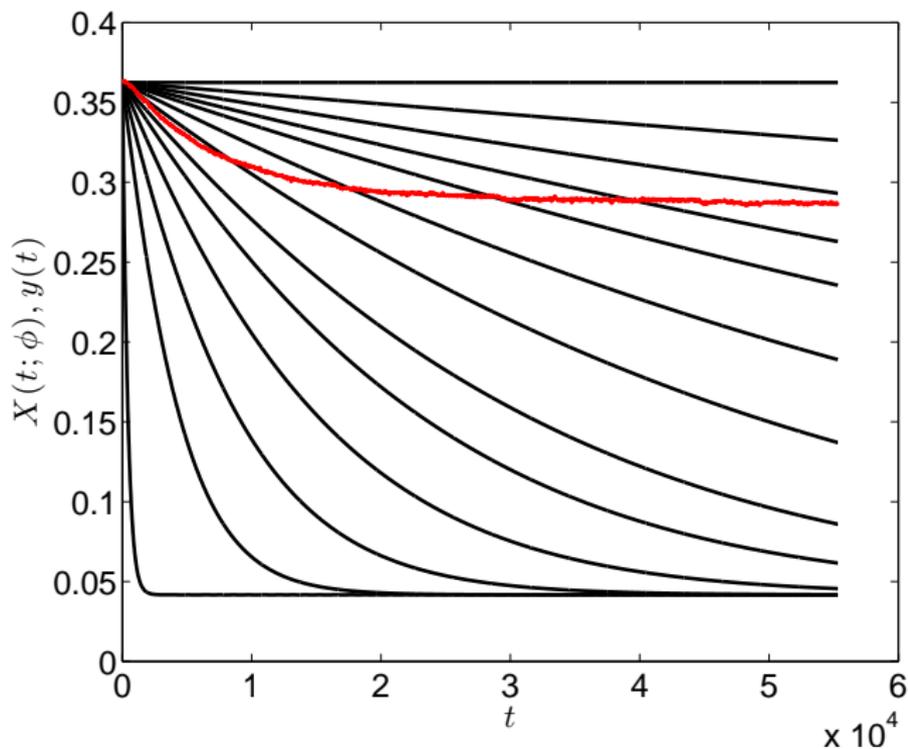
$$\frac{dX}{dt}(t) = I_0(1 - 10^{-X(t)})\phi_{\text{trans} \rightarrow \text{cis}}(\epsilon_{\text{trans}} - r_\phi \epsilon_{\text{cis}}) \left[\frac{r_\epsilon(r_\phi + 1)}{1 + r_\epsilon r_\phi} \frac{X(0)}{X(t)} - 1 \right]$$

ただし

- $r_\epsilon = \epsilon_{\text{cis}}/\epsilon_{\text{trans}}$
- $r_\phi = \phi_{\text{cis} \rightarrow \text{trans}}/\phi_{\text{trans} \rightarrow \text{cis}}$: 定数 (らしい)

以下, $\phi_{\text{trans} \rightarrow \text{cis}} = \phi$ のときの微分方程式の解を $X(t; \phi)$

$$F\left(t, X(t; \phi), \frac{dX}{dt}(t; \phi); \phi\right) = 0$$

量子収率 ϕ のときの吸光度 $X(t; \phi)$ と測定データ $y(t)$ 

発表の流れ

- 1 準備: 凸 2 次最適化問題とその解法
- 2 導入: フォトクロミック分子と量子収率分布
- 3 測定データと反応速度式
- 4 量子収率分布推定のための最適化モデル
- 5 数値実験

重ね合わせモデル

離散分布

仮定: 吸光度 = 各量子収率の分子の吸光度の重み付き和 + ノイズ

時刻 t における吸光度

量子収率 ϕ_i の分子の割合

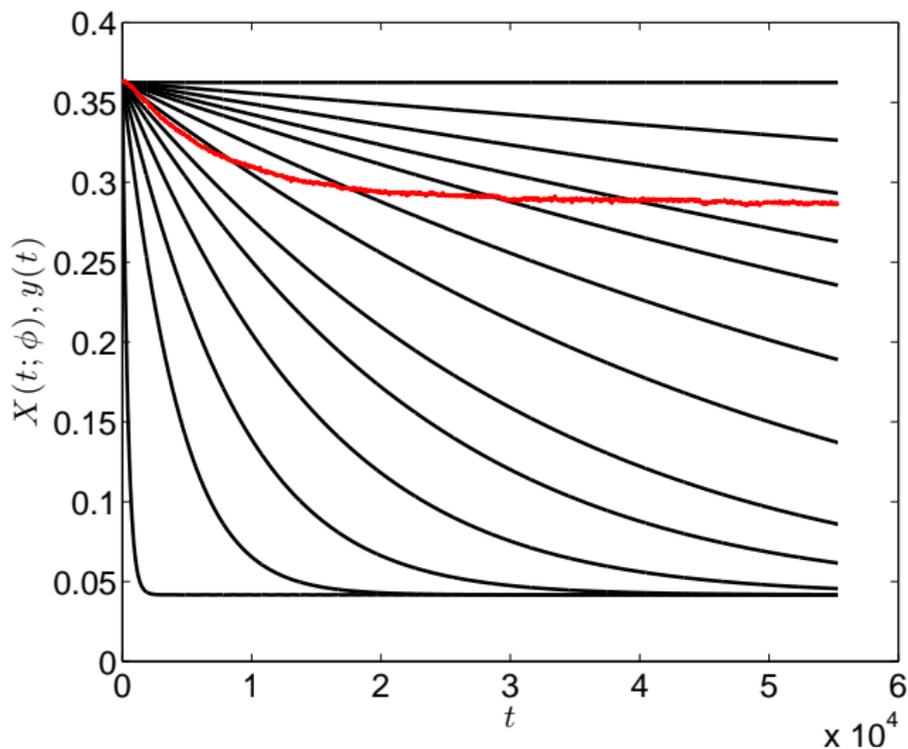
$$y(t) = \sum_{i=0}^N X(t; \phi_i) p_i + \text{ノイズ}$$

すべての分子の量子収率が ϕ_i と仮定したときの時刻 t における吸光度

連続分布

量子収率分布の密度関数 $\beta(\phi)$ on $0 \leq \phi \leq 1$

$$y(t) = \int_0^1 X(t; \phi) \beta(\phi) d\phi + \text{ノイズ}$$

量子収率 ϕ のときの吸光度 $X(t; \phi)$ と測定データ $y(t)$ 

基本モデル

- 目的関数: 測定値 $y(t)$ と理論値 $\int_0^1 X(t; \phi)\beta(\phi)d\phi$ の差の 2 乗和
- 制約条件: 反応速度式 + 密度関数の定義

$$\begin{array}{l}
 \text{minimize} \quad \sum_{t \in \mathcal{T}} \left(y(t) - \int_0^1 X(t; \phi)\beta(\phi)d\phi \right)^2 \\
 \text{subject to} \quad F \left(t, X(t; \phi), \frac{dX}{dt}(t; \phi); \phi \right) = 0 \quad (t \geq 0, 0 \leq \phi \leq 1) \\
 \int_0^1 \beta(\phi)d\phi = 1 \\
 \beta(\phi) \geq 0 \quad (0 \leq \phi \leq 1)
 \end{array}$$

無限次元の (常微分方程式制約付き) 最適化問題 ⇨ つらい

離散化モデル := 基本モデル + 離散化

- 量子収率 ϕ を $\phi_0 < \dots < \phi_N$ で離散化
- 積分を和分で近似, e.g., $\int_0^1 \beta(\phi) d\phi \approx \sum_{i=0}^{N-1} w_i \beta(\phi_i)$
- $F\left(t, X(t; \phi_i), \frac{dX}{dt}(t; \phi_i); \phi_i\right) = 0$ を数値的に解く for $i = 0, \dots, N$
 $\Rightarrow X(t; \phi_i)$ for $t \in \mathcal{T}$ をあらかじめ計算しておく

$$\left| \begin{array}{l} \text{minimize} \\ \text{subject to} \end{array} \right. \sum_{t \in \mathcal{T}} \left(y(t) - \sum_{i=0}^{N-1} X(t; \phi_i) \beta(\phi_i) w_i \right)^2$$

$$\sum_{i=0}^{N-1} \beta(\phi_i) w_i = 1$$

$$\beta(\phi_i) \geq 0 \quad (i = 0, \dots, N)$$

ただの有限次元の凸 2 次最適化問題 \Rightarrow これならかんたん (but ...)

正則化の必要性

細かい離散化 \Rightarrow 回帰モデルでいう “多重共線性” が起きる

$$\phi \approx \phi' \Rightarrow X(t; \phi) \approx X(t; \phi')$$

i.e., 目的関数の 2 次の係数に対応する行列は悪条件 \Rightarrow 結果が不安定

\Rightarrow “適当な” 正則化

- ~~L_1 正則化~~ \Leftarrow 疎性は期待しない
- L_2 正則化
- 曲率正則化

L_2 正則化モデル := L_2 正則化 + 離散化

$$\begin{array}{l}
 \text{minimize} \quad \sum_{t \in \mathcal{T}} \left(y(t) - \int_0^1 X(t; \phi) \beta(\phi) d\phi \right)^2 + \lambda \int_0^1 \beta(\phi)^2 d\phi \\
 \text{subject to} \quad F \left(t, X(t; \phi), \frac{dX}{dt}(t; \phi); \phi \right) = 0 \quad (t \geq 0, 0 \leq \phi \leq 1) \\
 \int_0^1 \beta(\phi) d\phi = 1 \\
 \beta(\phi) \geq 0 \quad (0 \leq \phi \leq 1)
 \end{array}$$

- リッジ回帰に制約がついたようなもの
- 正則化パラメータ λ を適切に選択する必要あり (λ : 大 \Rightarrow 正則化: 強)
- 離散化すると凸 2 次最適化問題

曲率正則化モデル := 導関数の Lipschitz 連続性 + 離散化

$$\begin{array}{l}
 \text{minimize} \quad \sum_{t \in \mathcal{T}} \left(y(t) - \int_0^1 X(t; \phi) \beta(\phi) d\phi \right)^2 \\
 \text{subject to} \quad F \left(t, X(t; \phi), \frac{dX}{dt}(t; \phi); \phi \right) = 0 \quad (t \geq 0, 0 \leq \phi \leq 1) \\
 \int_0^1 \beta(\phi) d\phi = 1 \\
 \beta(\phi) \geq 0 \quad (0 \leq \phi \leq 1) \\
 \left| \frac{d\beta}{d\phi}(\phi) - \frac{d\beta}{d\phi}(\phi') \right| \leq L |\phi - \phi'| \quad (0 \leq \phi', \phi \leq 1)
 \end{array}$$

- $\beta(\phi)$ の曲率を制限したようなもの
- 正則化パラメータ L を適切に選択する必要あり (L : 小 \rightarrow 正則化: 強)
- 離散化すると凸 2 次最適化問題 (微分は差分で近似)

発表の流れ

- 1 準備: 凸 2 次最適化問題とその解法
- 2 導入: フォトクロミック分子と量子収率分布
- 3 測定データと反応速度式
- 4 量子収率分布推定のための最適化モデル
- 5 数値実験

数値実験 1: 既知の量子収率分布を正しく計算できるか?

手順

1. 真の量子収率分布 $\beta(\phi)$ を適当に決める
2. $F\left(t, X(t; \phi), \frac{dX}{dt}(t; \phi); \phi\right) = 0$ を数値的に解く for $\phi = \phi_0, \dots, \phi_N$
(ただし $N = 1000$ として一様に離散化)
3. $y(t) \approx \int_0^1 X(t; \phi)\beta(\phi)d\phi$ を台形公式で計算 for $t \in \mathcal{T}$
 ⇨ これを仮想的な測定値とする
4. L_2 正則化モデルと曲率正則化モデルを解く
 ⇨ 最適解 $\hat{\beta}$ を真の分布 β との差を調べる

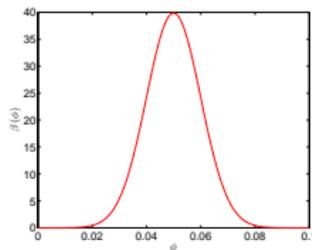
環境

- Mac OS X 10.9.5, 2.3 GHz Intel Core i7, 8GB
- 微分方程式: ode45 in MATLAB R2014a ((4, 5) 次の Runge-Kutta)
- 凸 2 次最適化問題: Gurobi 5.6.3
- 正則化パラメータ: **いろいろ試してみる**

正解として用いた量子収率分布

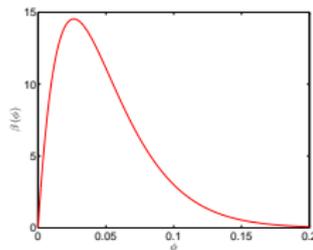
正規分布

- $N(0.05, 0.01^2)$
- $N(0.05, 0.005^2)$



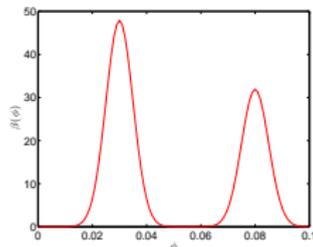
ベータ分布

- $Be(1, 19)$
- $Be(2, 38)$



混合正規分布

- $Mix_A = 0.25N(0.02, 0.005^2) + 0.75N(0.06, 0.005^2)$
- $Mix_B = 0.6N(0.03, 0.005^2) + 0.4N(0.08, 0.005^2)$



真の分布と最適解との差 with L_2 正則化モデル

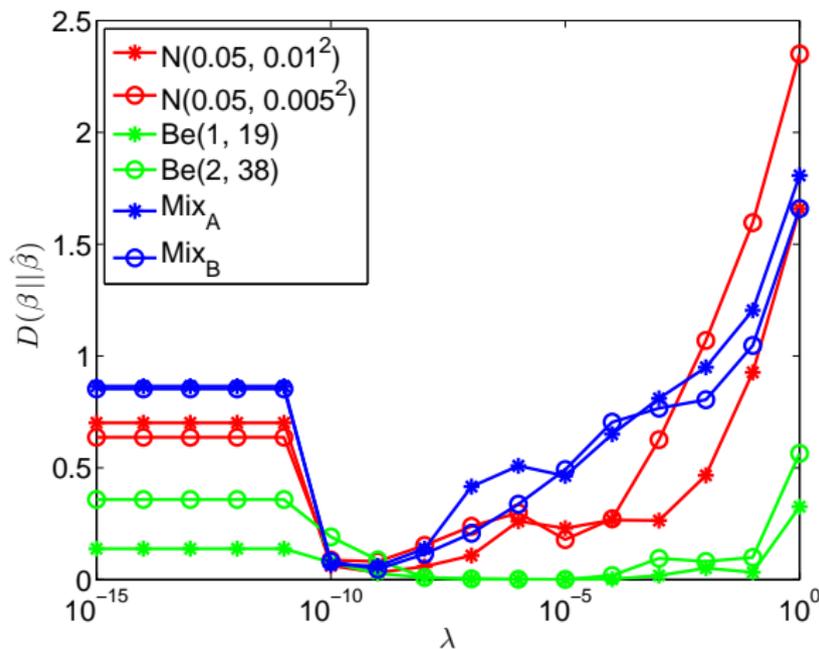
真の分布 β の最適解 $\hat{\beta}$ に対する Kullback-Leibler 情報量

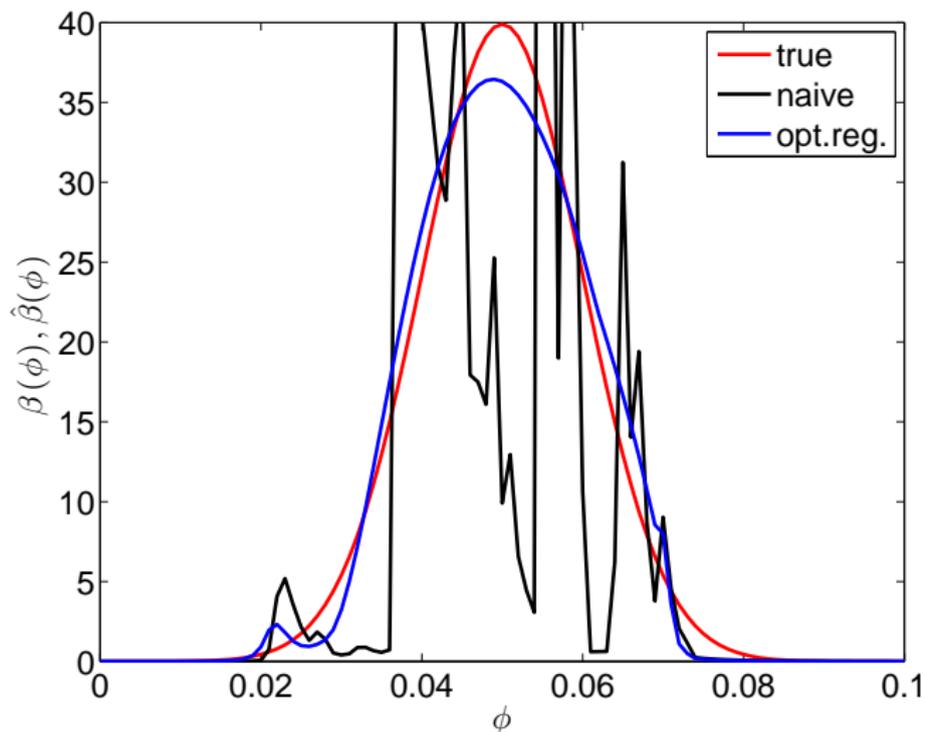
$$D(\beta||\hat{\beta}) = \int_0^1 \beta(\phi) \log \frac{\beta(\phi)}{\hat{\beta}(\phi)} d\phi$$

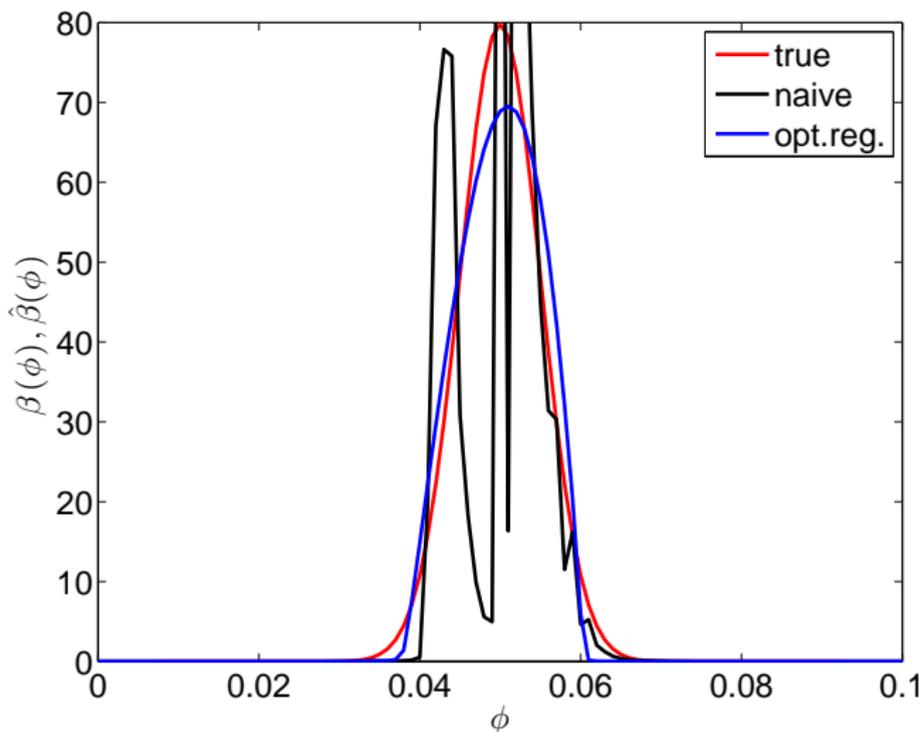
λ	$N(.05, .01^2)$	$N(.05, .005^2)$	$Be(1, 19)$	$Be(2, 38)$	Mix_A	Mix_B
0	7.0e-01	6.4e-01	1.4e-01	6.2e-01	8.6e-01	8.5e-01
1e-10	6.2e-02	8.7e-02	7.6e-02	4.0e-01	6.7e-02	8.2e-02
1e-09	3.2e-02	8.1e-02	2.7e-02	3.1e-01	6.0e-02	4.6e-02
1e-08	5.8e-02	1.5e-01	9.3e-03	2.3e-01	1.4e-01	1.1e-01
1e-07	1.1e-01	2.4e-01	3.1e-03	2.1e-01	4.2e-01	2.1e-01
1e-06	2.6e-01	3.0e-01	1.5e-03	2.1e-01	5.1e-01	3.4e-01
1e-05	2.3e-01	1.8e-01	4.1e-04	2.0e-01	4.6e-01	4.9e-01
1e-04	2.7e-01	2.7e-01	2.1e-03	3.1e-01	6.5e-01	7.0e-01
1e-03	2.6e-01	6.3e-01	1.9e-02	7.9e-01	8.1e-01	7.7e-01
1e-02	4.7e-01	1.1e+00	5.1e-02	7.3e-01	9.5e-01	8.0e-01
1e-01	9.3e-01	1.6e+00	3.3e-02	1.1e-01	1.2e+00	1.0e+00

真の分布と最適解との差 with L_2 正則化モデル

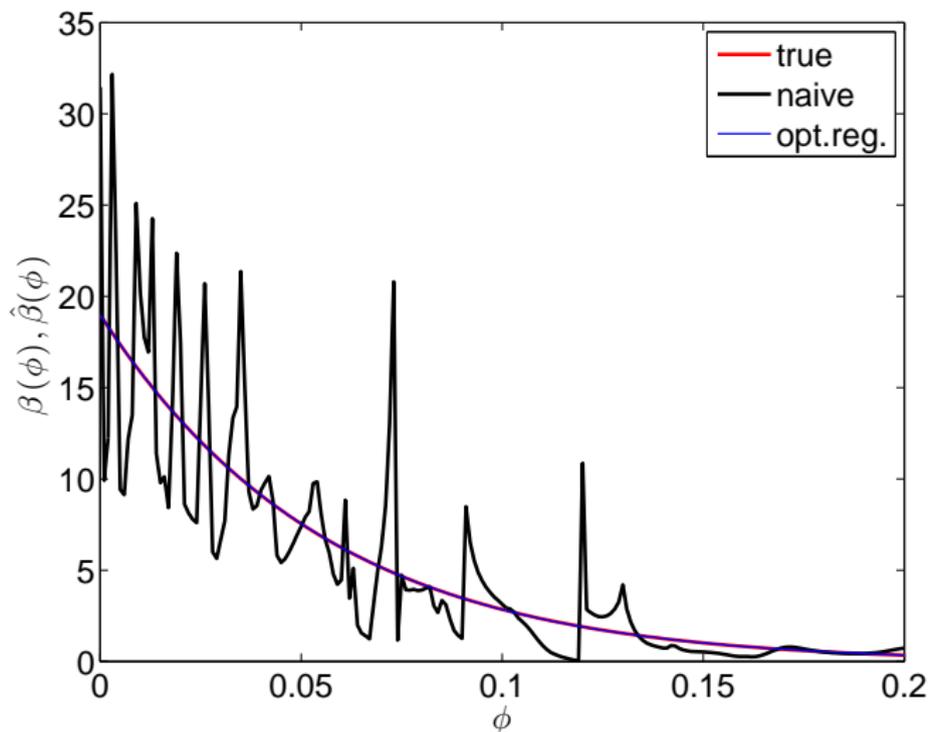
正解 β の最適解 $\hat{\beta}$ に対する KL 情報量 $D(\beta||\hat{\beta}) = \int_0^1 \beta(\phi) \log \frac{\beta(\phi)}{\hat{\beta}(\phi)} d\phi$



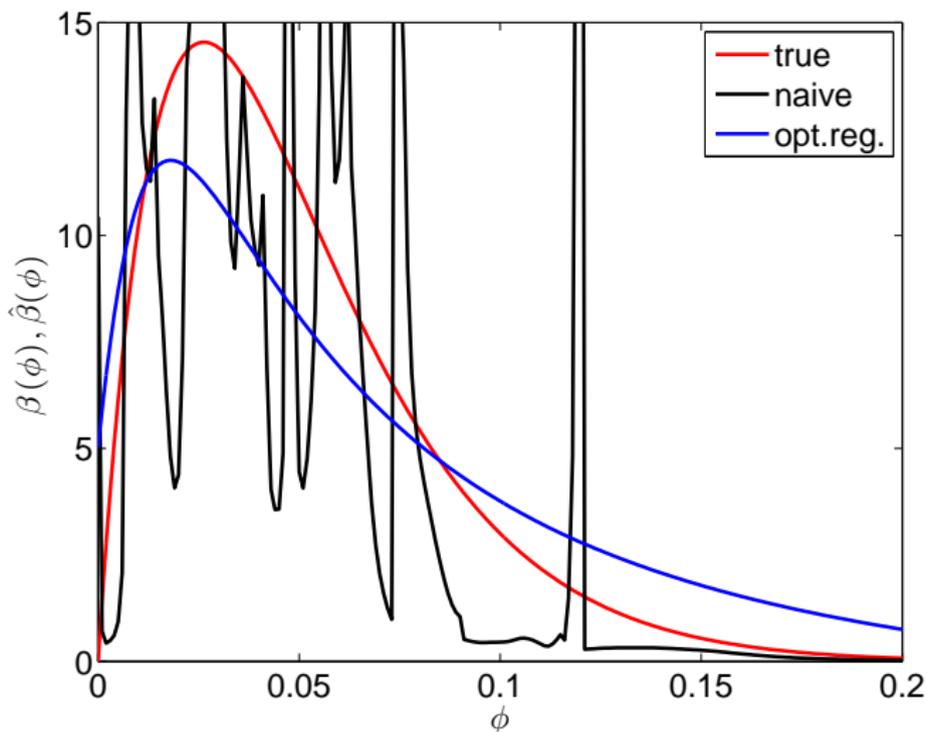
結果: $N(0.05, 0.01^2)$ with L_2 正則化モデル

結果: $N(0.05, 0.005^2)$ with L_2 正則化モデル

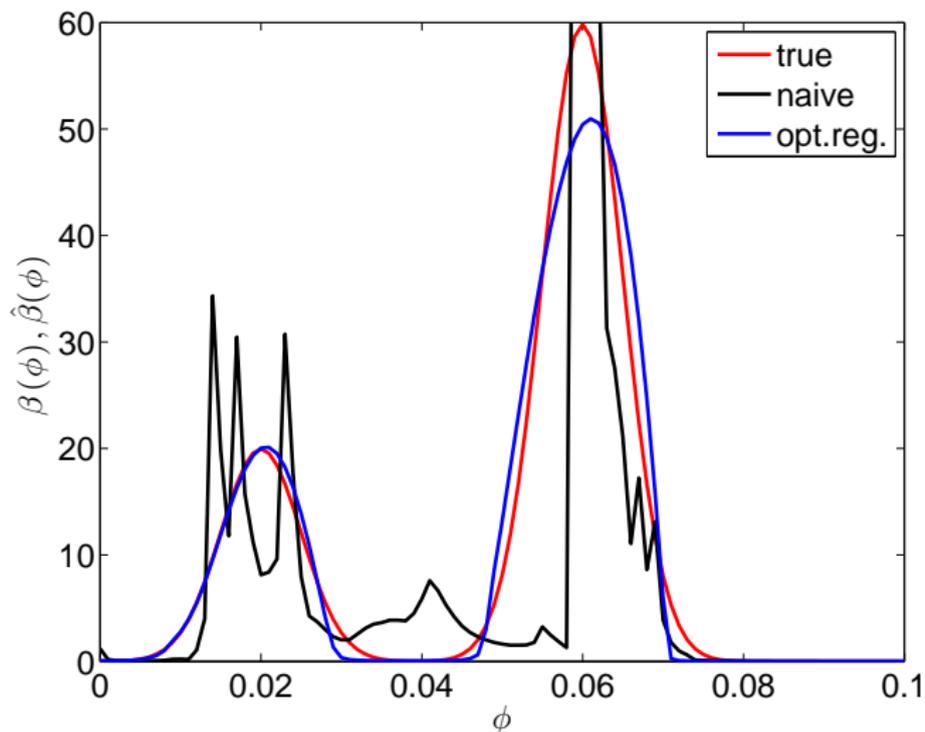
結果: Be(1, 19) with L_2 正則化モデル

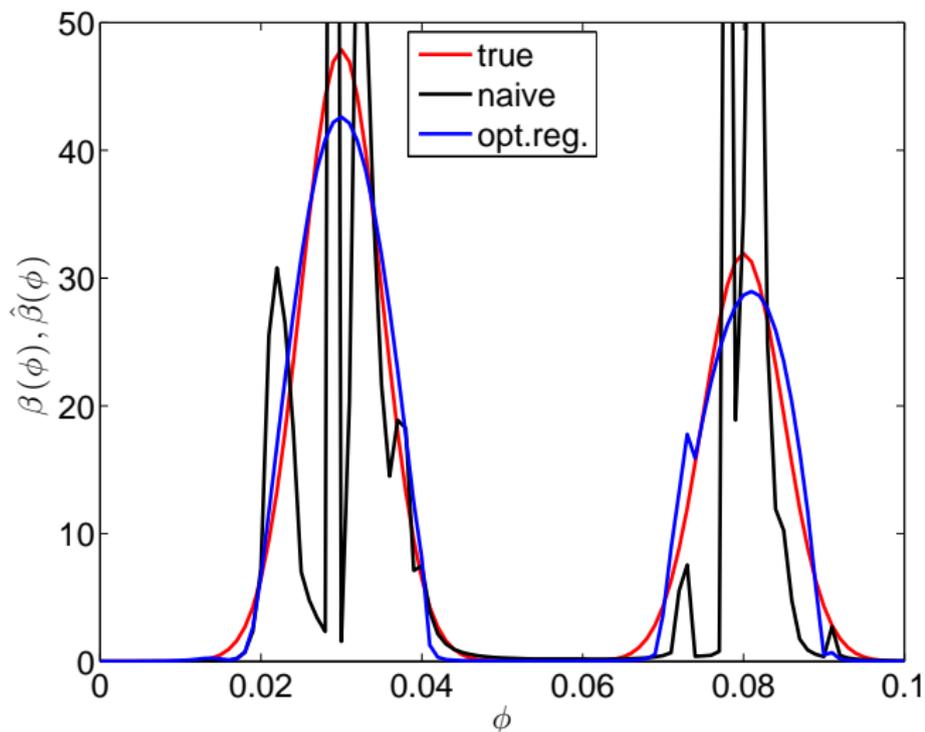


結果: Be(2, 38) with L_2 正則化モデル



結果: Mix_A with L_2 正則化モデル



結果: Mix_B with L_2 正則化モデル

真の分布と最適解との差 with 曲率正則化モデル

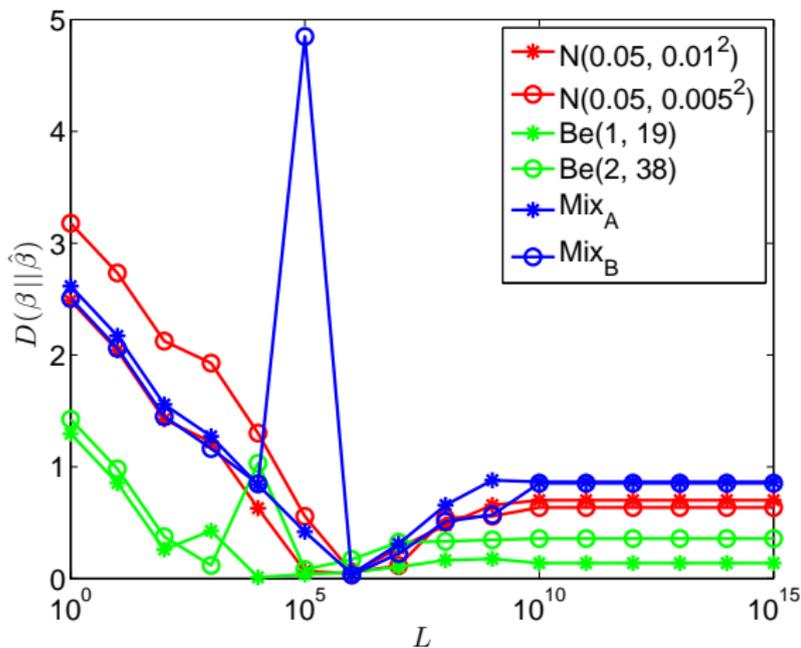
真の分布 β の最適解 $\hat{\beta}$ に対する Kullback-Leibler 情報量

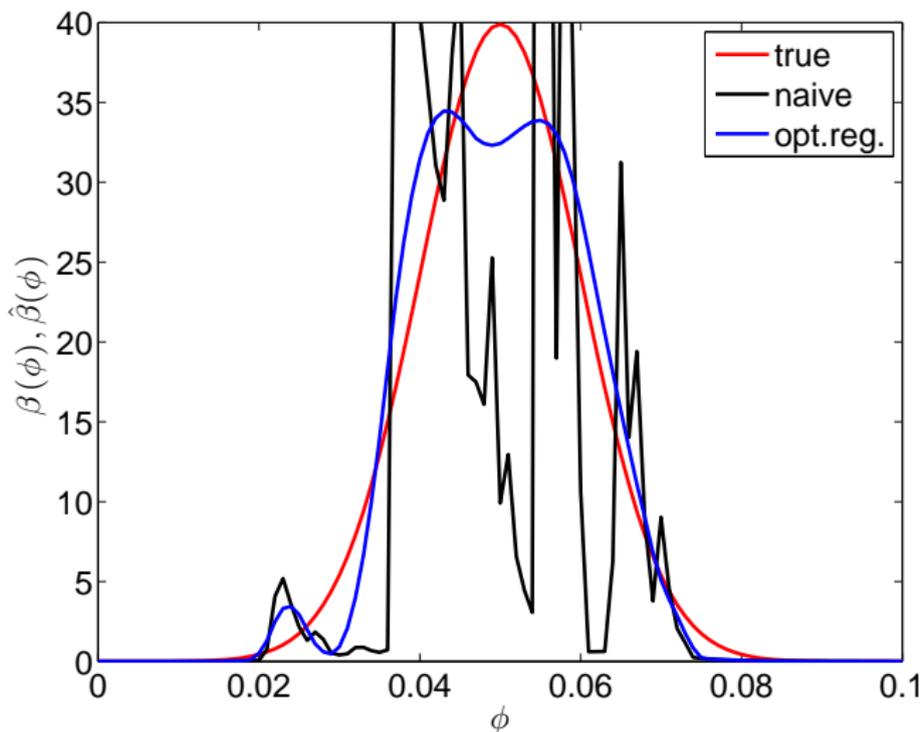
$$D(\beta||\hat{\beta}) = \int_0^1 \beta(\phi) \log \frac{\beta(\phi)}{\hat{\beta}(\phi)} d\phi$$

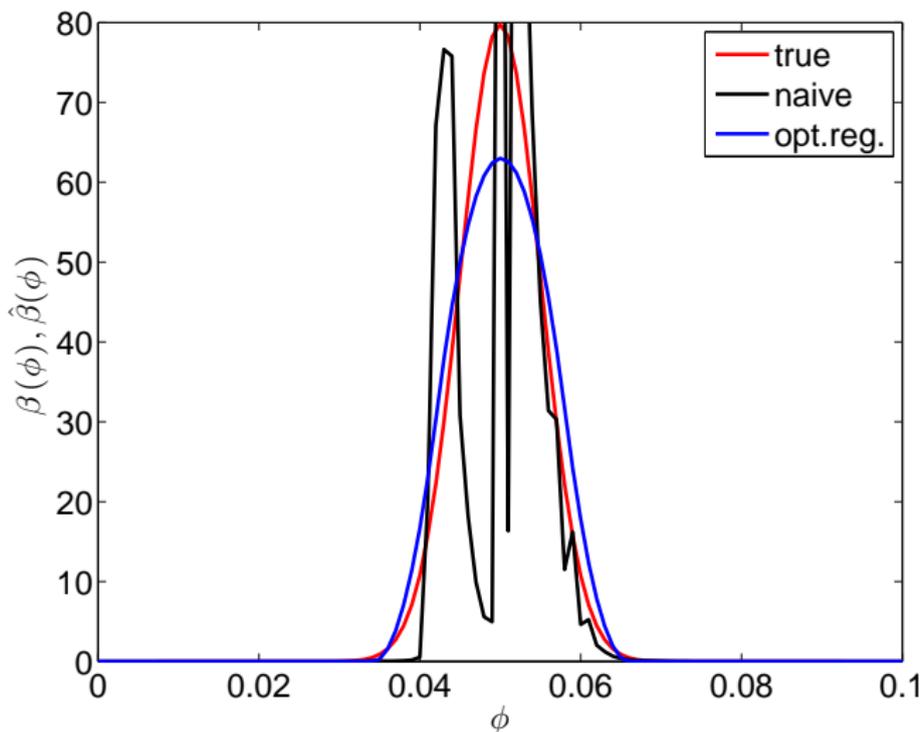
L	$N(.05, .01^2)$	$N(.05, .005^2)$	$Be(1, 19)$	$Be(2, 38)$	Mix_A	Mix_B
1e+10	7.0e-01	6.4e-01	1.4e-01	3.6e-01	8.6e-01	8.5e-01
1e+09	6.6e-01	5.6e-01	1.8e-01	3.5e-01	8.8e-01	5.7e-01
1e+08	4.8e-01	5.4e-01	1.7e-01	3.3e-01	6.6e-01	5.1e-01
1e+07	2.7e-01	1.2e-01	1.0e-01	3.3e-01	3.1e-01	2.2e-01
1e+06	4.7e-02	5.7e-02	5.7e-02	1.7e-01	4.0e-02	3.2e-02
1e+05	6.6e-02	5.6e-01	3.8e-02	8.2e-02	4.2e-01	4.8e+00
1e+04	6.3e-01	1.3e+00	1.1e-02	1.0e+00	8.4e-01	8.5e-01
1e+03	1.2e+00	1.9e+00	4.3e-01	1.2e-01	1.3e+00	1.2e+00
1e+02	1.4e+00	2.1e+00	2.6e-01	3.7e-01	1.6e+00	1.4e+00
1e+01	2.0e+00	2.7e+00	8.5e-01	9.8e-01	2.2e+00	2.1e+00
1e+00	2.5e+00	3.2e+00	1.3e+00	1.4e+00	2.6e+00	2.5e+00

真の分布と最適解との差 with 曲率正則化モデル

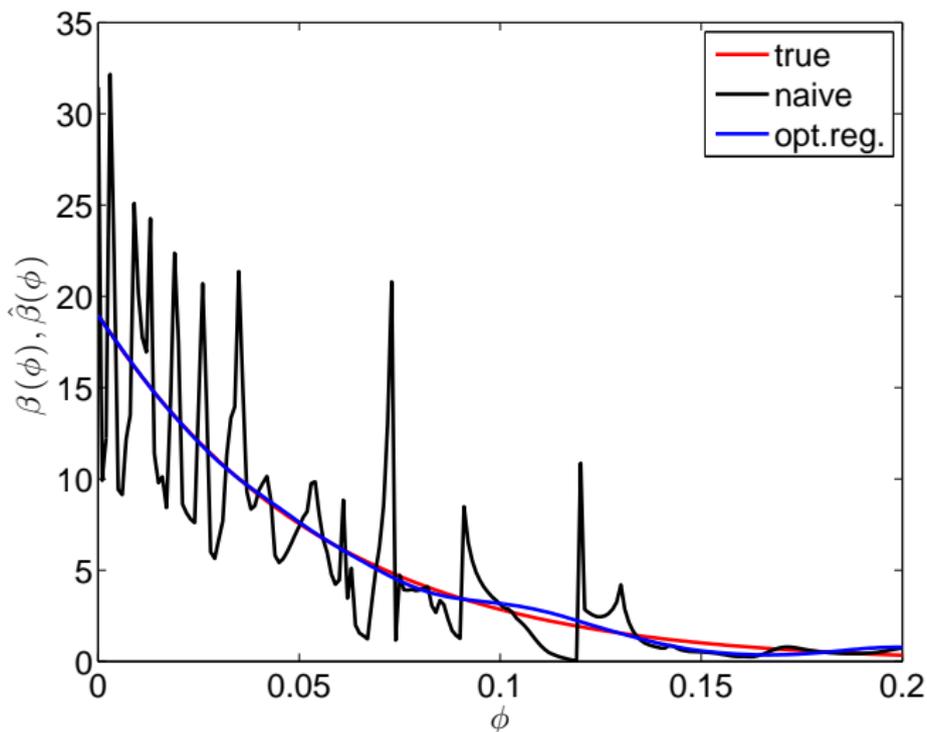
正解 β の最適解 $\hat{\beta}$ に対する KL 情報量 $D(\beta||\hat{\beta}) = \int_0^1 \beta(\phi) \log \frac{\beta(\phi)}{\hat{\beta}(\phi)} d\phi$



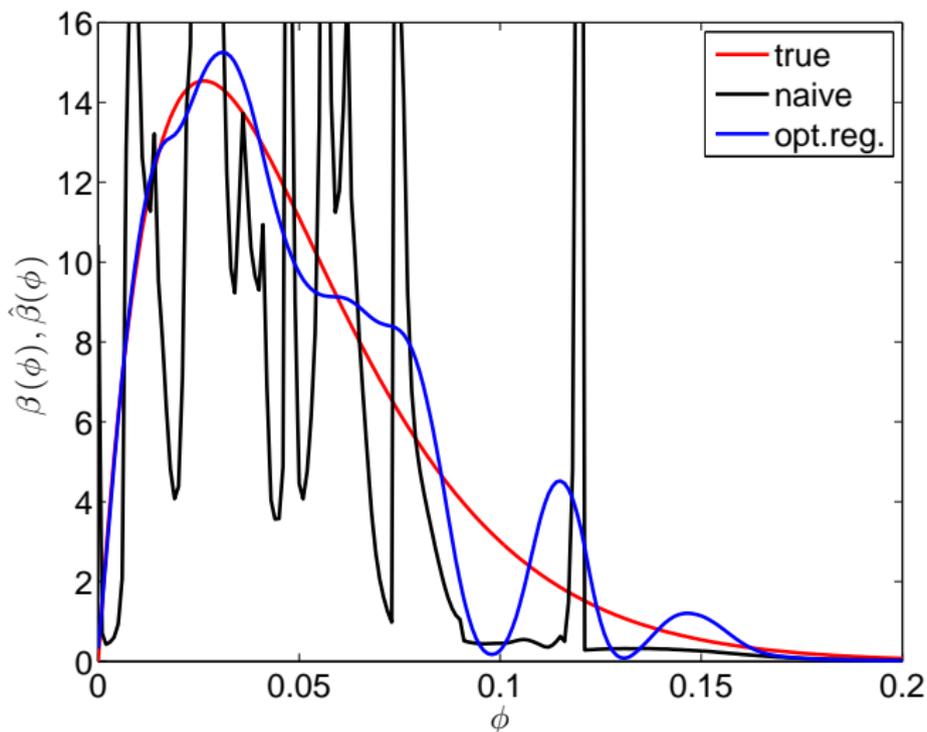
結果: $N(0.05, 0.01^2)$ with 曲率正則化モデル

結果: $N(0.05, 0.005^2)$ with 曲率正則化モデル

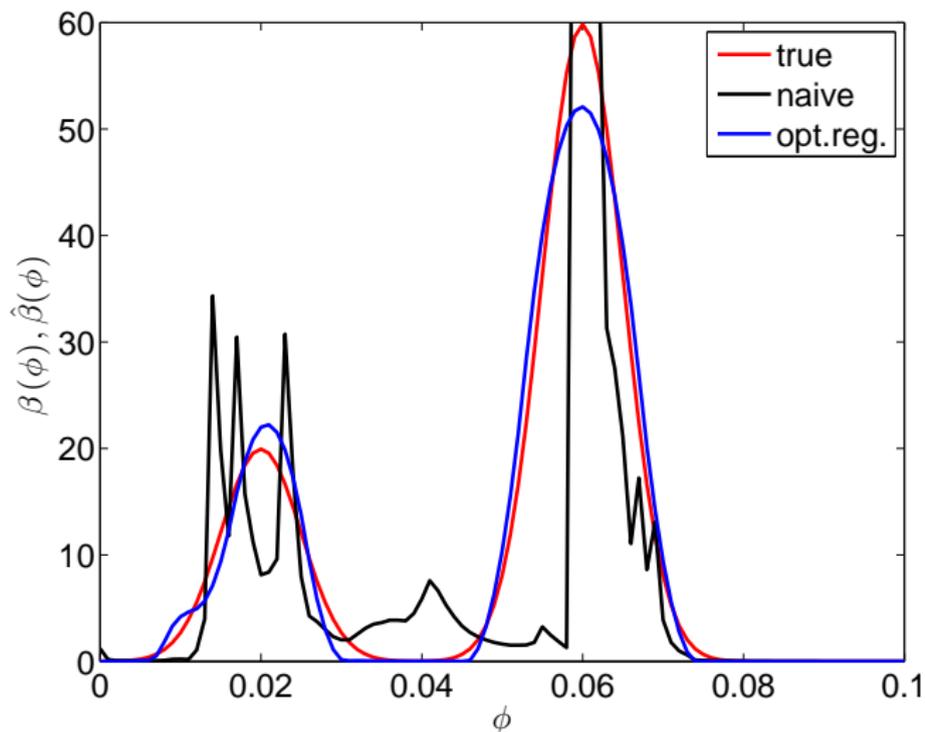
結果: Be(1, 19) with 曲率正則化モデル

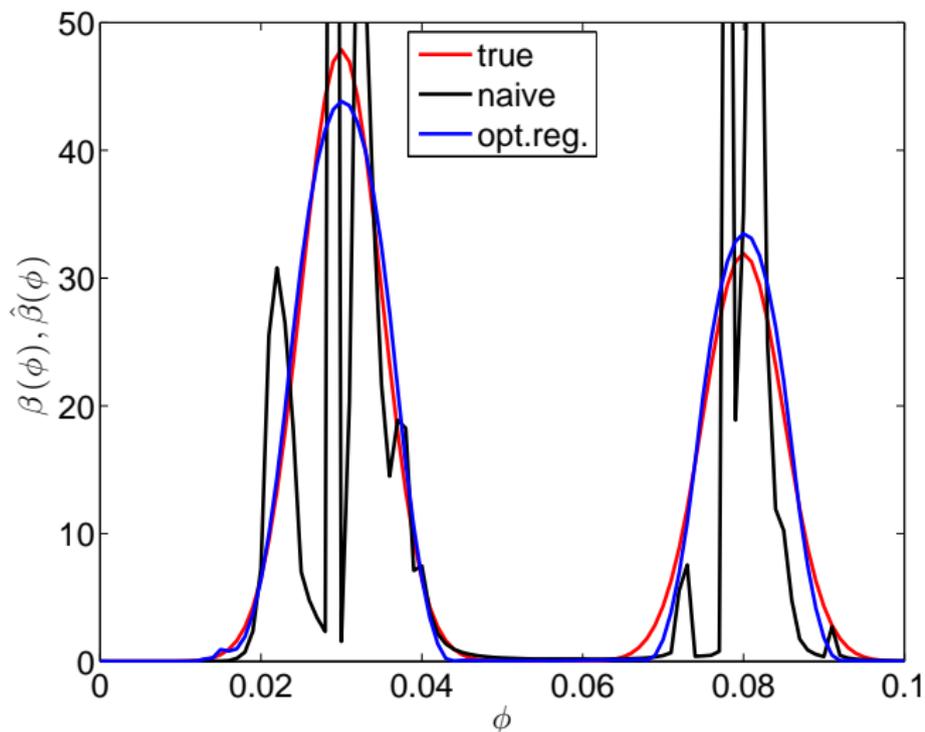


結果: Be(2, 38) with 曲率正則化モデル



結果: Mix_A with 曲率正則化モデル



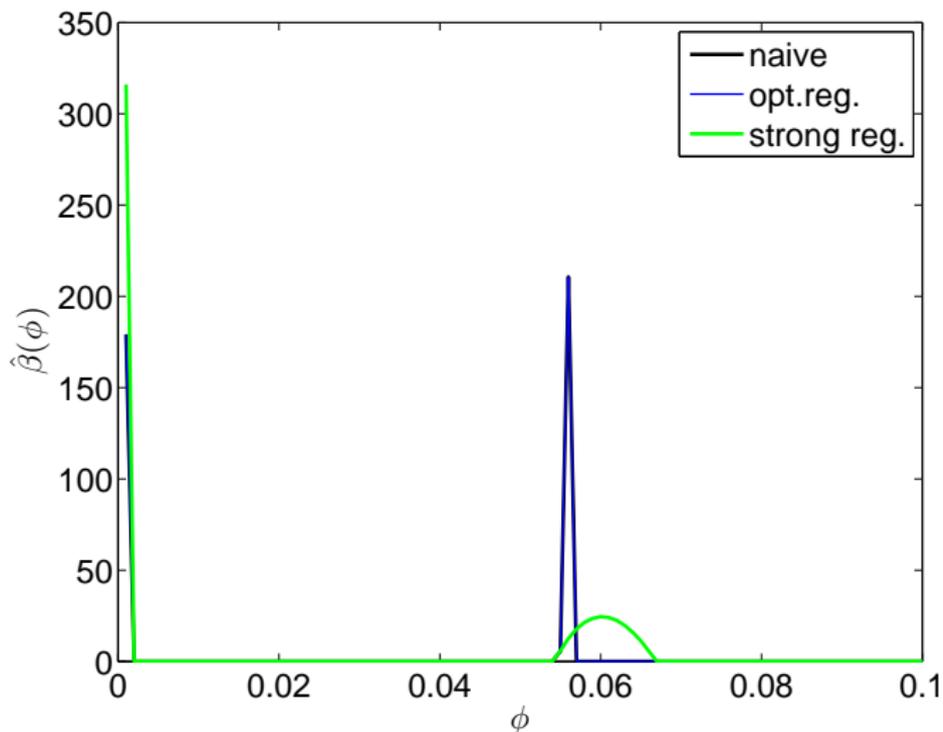
結果: Mix_B with 曲率正則化モデル

数値実験 2: 実際の測定値を用いた計算

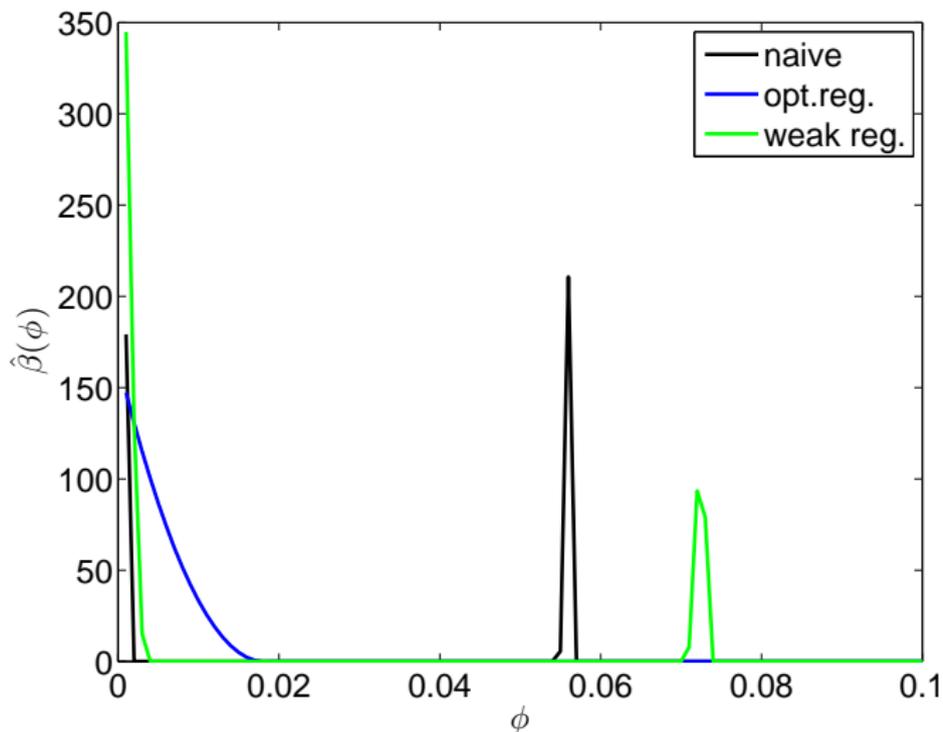
実際の測定値を用いて L_2 正則化モデルと曲率正則化モデルを解いてみる

環境

- Mac OS X 10.9.5, 2.3 GHz Intel Core i7, 8GB
- 微分方程式: ode45 in MATLAB R2014a ((4, 5) 次の Runge-Kutta)
- 凸 2 次最適化問題: Gurobi 5.6.3
- 正則化パラメータ: 実験 1 で結果の良かったもの + いくつか

結果: 測定値 with L_2 正則化モデル

結果: 測定値 with 曲率正則化モデル



まとめ・今後の課題

まとめ

- 目的: 高分子固体中のフォトクロミック分子の量子収率分布の推定
- モデル: 重ね合わせモデル
- 定式化: 関数空間における常微分方程式制約付き最適化問題
 - $\xrightarrow{\text{離散化}}$ 凸 2 次最適化問題
 - $\xrightarrow{\text{正則化}}$ ノイズに対して頑健
- 結果: 人工データに対してはますます & 分布の概形は把握可能

今後の課題

- 正則化パラメータの適切な値の決め方?
- もうちょっと分布らしくなる測定データで計算を...