

## 研究の内容

### 研究の動機

- 「情報」の数理的性質を明らかにしたい。
- 実社会の情報処理に役立つ技術を開発したい。

そのために以下のようなテーマに取り組んでいます。

**機械学習と最適化**：多様な構造のデータから有用な情報を取り出すための、効率的な学習アルゴリズムの開発

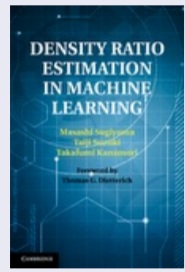
**数理統計学**：データ解析を適切に行うための統計手法の開発。不確実な状況下でのリスク尺度や損失関数の設計に関する研究。

**情報幾何学**：「情報科学における一般相対論」。双対構造（外延と内包）と不変性に関する幾何学の研究。数理科学全般への応用。

## 機械学習と最適化

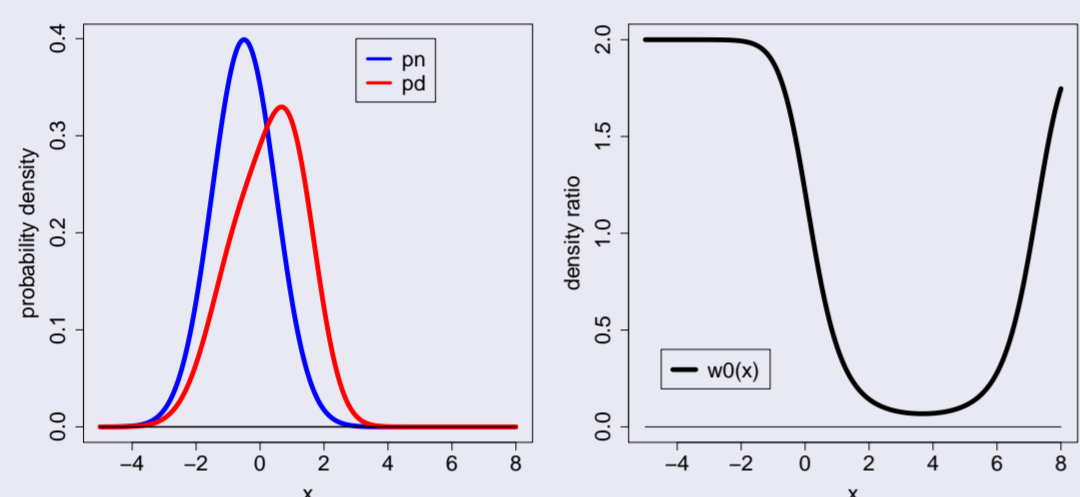
### 密度比推定とその応用

密度比に関する研究成果を出版



M. Sugiyama, T. Suzuki, T. Kanamori,  
Density Ratio Estimation in Machine Learning,  
Cambridge University Press, 2012.

密度比とは：確率密度  $p_n(x)$ ,  $p_d(x)$  の比  $w_0(x) = \frac{p_n(x)}{p_d(x)}$



密度比の重要性：さまざまな応用

- 共変量シフト下での回帰分析。
- 分布間のダイバージェンスや距離の推定：2標本検定，次元削減など。

密度比の推定法

$p_n(x)$ ,  $p_d(x)$  をバラバラに推定して比を取る方法  
⇒ 非常に不安定。

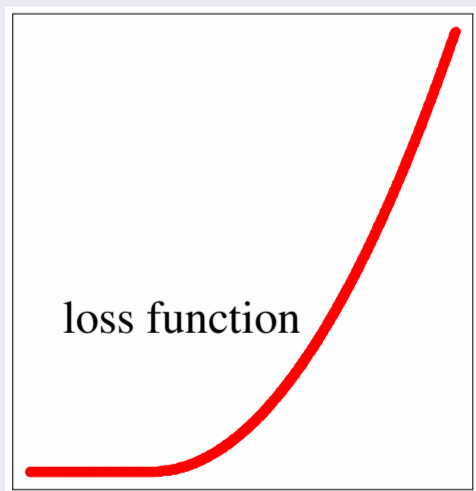
■ 密度比  $w_0(x)$  を高精度に直接推定する方法を提案

- カーネル化した推定量の統計的性質を経験過程の理論で解析
- 計算効率を smoothed analysis の理論で解析
- 統計解析言語 R による実装。snow library による並列化

### 機械学習における共役性

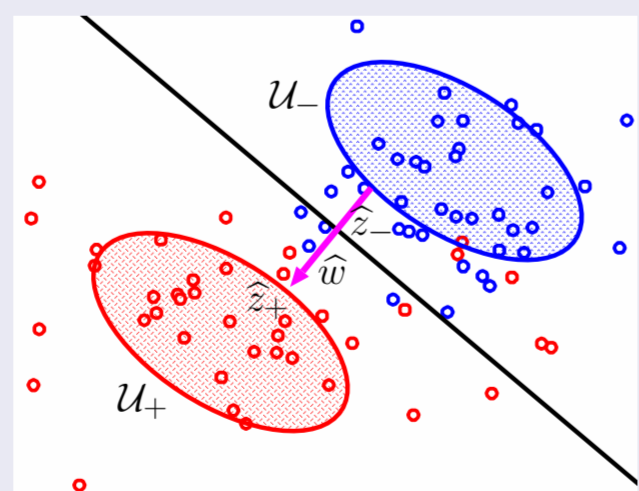
パターン認識の2つのアプローチ：共役な関係がある。

#### 損失関数アプローチ



統計的性質  
直感的に分りにくい  
(モデリングしにくい)

#### 不確実性集合アプローチ



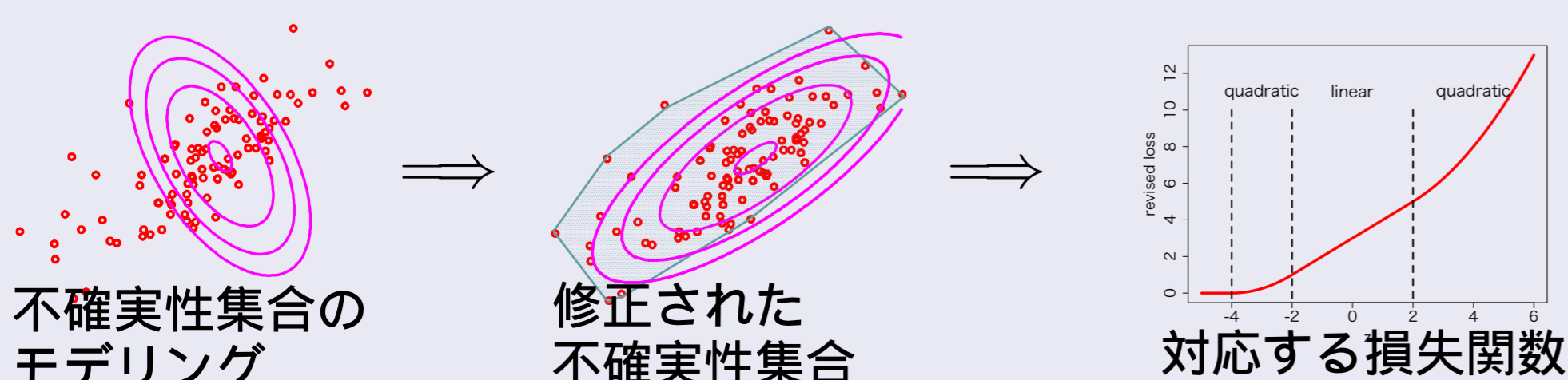
統計的性質  
直感的に分りにやすい  
(モデリングしやすい)

共役性

共役性を介して...

不確実性集合を，対応する損失関数が存在するように修正。

■ 直感的なモデリングと優れた統計的性質を両立。



[K, Takeda, Suzuki, '13]

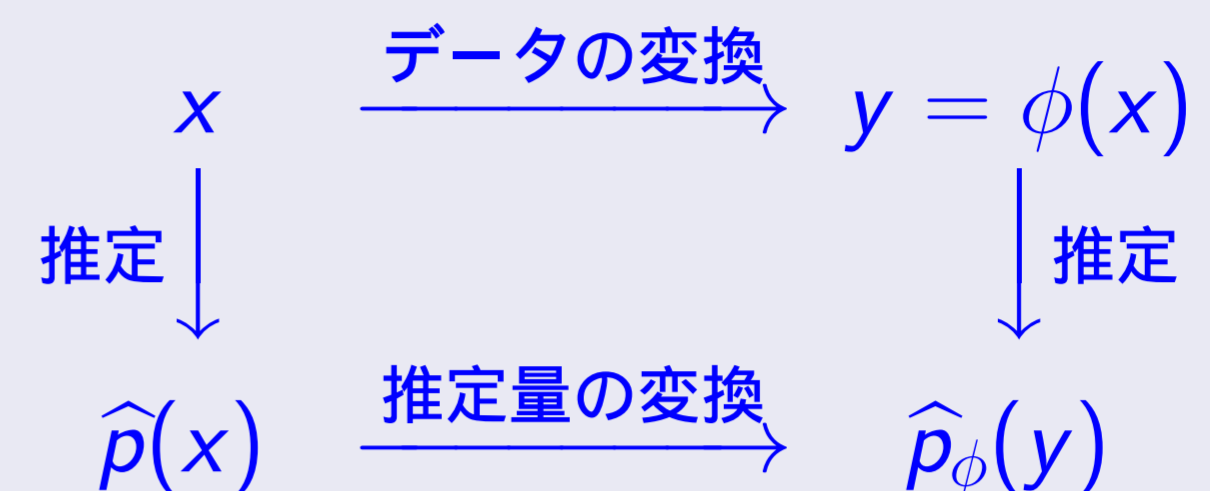
## 数理統計学

### 統計的推論と不変性

大切なこと：データ変換に対する不変性

例：メートルで測定したデータとセンチで測定したデータ。統計解析の結果は本質的に同じであってほしい。

不変な推定：以下の可換図式が成り立つ推定



- 任意のデータ変換に対して不変：最尤推定量のみ。
- アフィン変換に対して不変：[K&Fujisawa, '13]で考察。不変性とその他の統計的性質とのトレードオフを研究。

「データ変換の不変性」から既存の推定法を特徴付け。新しい推定量のクラスを提案：ロバスト性など良い性質。

## 情報幾何学

### 準ニュートン法の解析

長年謎だった問題を情報幾何的アプローチで解明

関数  $f(x)$  の最小化問題

$$\min_x f(x), \quad x \in \mathbb{R}^n.$$

準ニュートン法による反復アルゴリズム

$$x_{k+1} = x_k - B_k^{-1} \nabla f(x_k), \quad k = 0, 1, 2, \dots$$

$$\{x_k\}_{k=1}^{\infty} : (\text{局所}) \text{最適解へ収束}$$

正定値行列  $B_k$  の更新則によって，さまざまなアルゴリズムが提案されている。

■ BFGS更新則，DFP更新則，Broyden更新則...

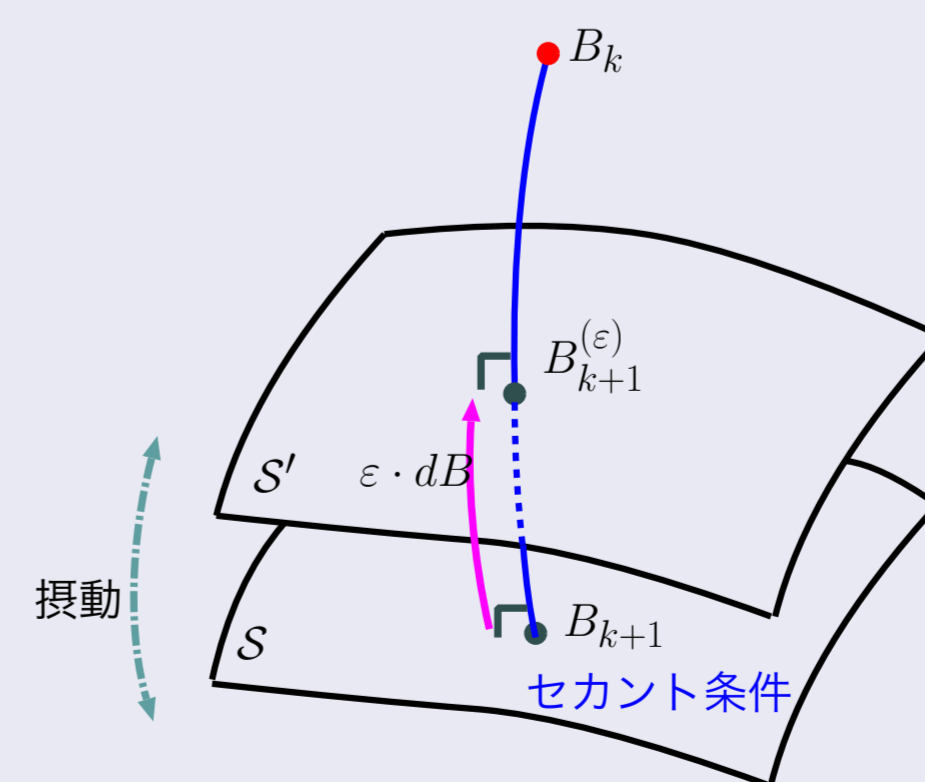
BFGS更新則が数値的に優れている。その理由は？  
長年よく分かっていなかった。

行列  $B_k$  の更新則：多変量正規分布の情報幾何と深く関連。

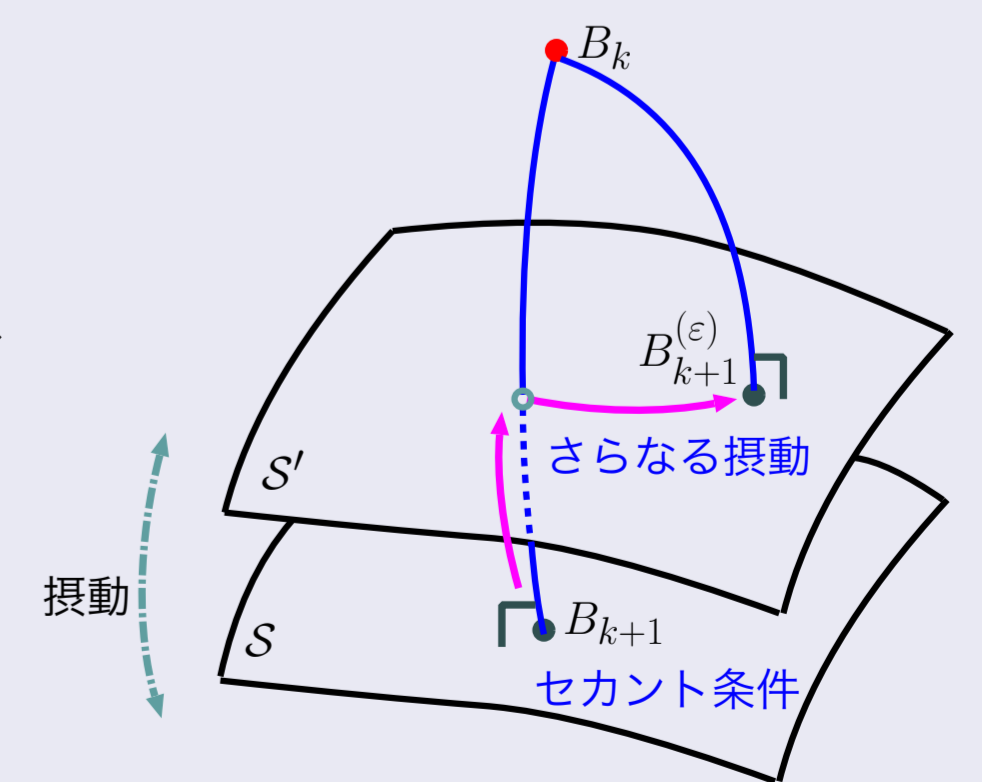
- 幾何学的直観&ロバスト(頑健)統計の手法  
⇒ アルゴリズムの安定性解析
- 直線探索の誤差に対して BFGS だけがロバスト (gross error 有限)。  
BFGS 以外はさらなる摂動が加わり不安定 (gross error 無限)。

準ニュートン法の幾何的描像 [K&Ohara, '13]

#### BFGS更新則



#### その他の更新則



他にも以下のようなさまざまな研究テーマに取り組んでいます：

ブースティング，半教師付き学習，ランキング学習，能動学習，分位点回帰，不確実性の下での最適化，ロバスト最適化，グループテスト，ベイジアンネットワーク，数理ファイナンスにおけるリスク尺度...

研究成果は JMLR, NIPS, ICML, COLT などのトップジャーナルやトップ国際会議で発表されています。

興味のある方，ぜひ一緒に研究しましょう！詳しくは以下を参照下さい。

<http://www.math.cm.is.nagoya-u.ac.jp/~kanamori>