EMアルゴリズム ：
クラスタリングへの適用と最近の発展

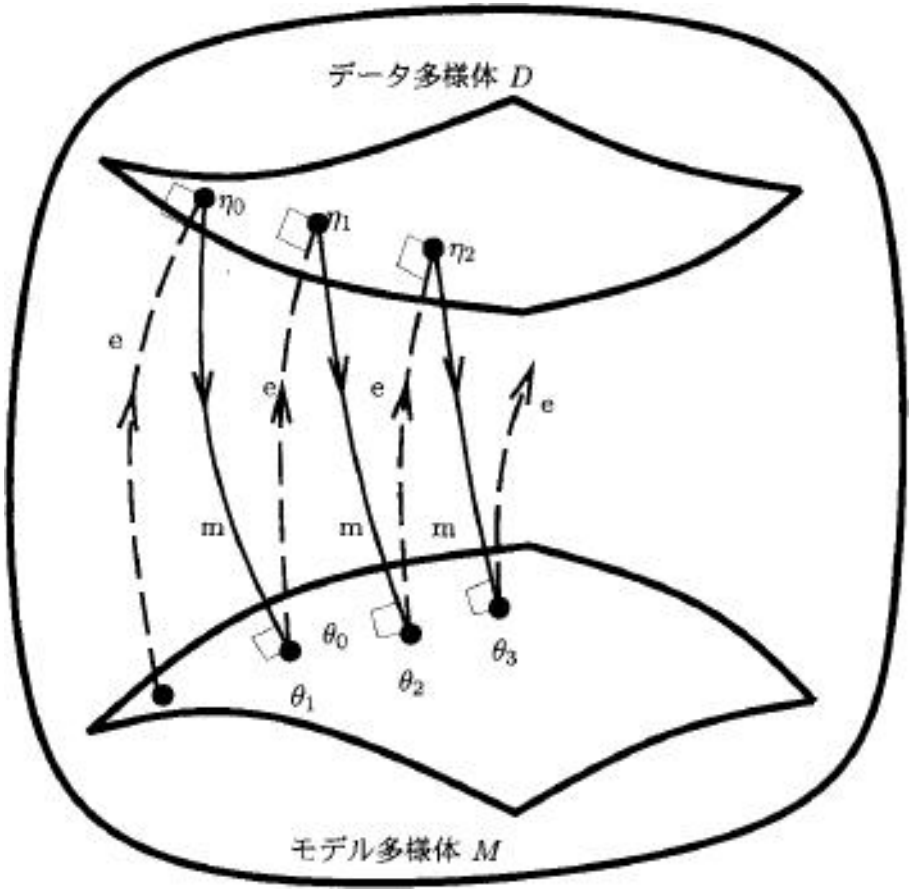
6. EMアルゴリズムの発展
7. まとめ

6.1　EMアルゴリズムの直観的イメージ(1/2)

- 情報幾何では統計的推定を
  確率分布の空間での操作としてとらえる
  - 通常の統計的推定では観測データが分布の空間の
    1点に対応づけられ、そこからモデルの空間への
    射影を取ったものが推定量となる
  - EMアルゴリズムで考えるような不完全な観測では
    観測データは完全変数の分布の空間の中では
    1点ではあらわせない
    →不完全な分の広がりを持った空間となる

6.1　EMアルゴリズムの直観的イメージ(2/2)

- データとモデルがともに空間をなしているのでその距離が最も近いところを推定量に採用
- →e(exponential)とm(mix-ture)の2つの射影を繰り返し最近点を見つける
- emとEMの二つのアルゴリズムは一般的な条件もとで一致する



EMアルゴリズムの幾何学的イメージ

6.2　ベイズ推定への拡張

- ベイズ推定とは
  - パラメータを確率変数としてその分布を考え、
    事後分布$p(\theta|x)$をできるだけ大きくしようとする
    パラメータを求めること
  - θに冠する事前分布$r(\theta)$を設定する必要がある
  - EMアルゴリズムのMステップにおいて$Q(\theta|\theta^{(t)})$の代わりに
    $Q(\theta|\theta^{(t)})+\log r(\theta)$
    を最大化する

6.3 初期化と局所最適解

- 局所最適解から脱する
  - ホモトピー法の導入
    - 単純な問題から次第に解くべき難しい問題へと変形していく方法で、決定論的アニーリングアルゴリズムと呼ばれている
- 混合分布の場合の局所最適解への収束を避ける
  - SMEMアルゴリズム
    - 一度収束したクラスタを分割したり併合したりするメカニズムを導入

6.4 高速化

6.4.1 2次収束アルゴリズム

- EMアルゴリズムは1次収束のアルゴリズムのため
  最適解の近傍での収束はかなり遅い
  →共役勾配法やAitken加速といった汎用的な
  方法により2次収束するようなアルゴリズムが
  多数提案されている
  - これらのアルゴリズムは
    1ステップあたりの計算量を増やしてしまう
    →全体の計算時間やインブリメンテーションの複雑度に
    応じて適当に選ぶ必要がある

6.4.2 Mステップの近似

- Mステップで$Q(\theta|\theta^{(t)})$を最大化することが難しい場合は制限を緩めて
  $Q(\theta|\theta^{(t)}) \geqq Q(\theta^{(t)}|\theta^{(t)})$
  を満たすθを見つける
  - この方法でも尤度の単調増加性は保たれている
  - これを一般化EMアルゴリズムと呼ぶ

6.4.3 Eステップの近似(1/2)

- モンテカルロシミュレーションによる近似計算
  - Eステップでの条件付の期待値を計算する際の計算量を減らす工夫
  - ただし、一般に乱数を生成する分布$s(x)$が複雑になることが多く、定常分布が$s(x)$になるようなマルコフ連鎖$\tilde{s}\left(x_{t+1}\middle|x_t\right)$を用いてモンテカルロ平均をとることがある
    - これをMCMC（マルコフ連鎖モンテカルロ）法と呼ぶ

6.4.3 Eステップの近似(2/2)

- 平均場近似
  - 一般にモンテカルロシミュレーションを行っても
    計算時間が長いことが多いのでモデルの単純化の
    仮定をおいて簡単な形にして問題を解く
  - 大胆な単純化は計算量を大幅に減らせるが
    その分最適解でないところに収束することもある
  - 計算量を減らしつつもモデルの構造を
    大きく壊さない近似法が求められている

7. まとめ

- EMアルゴリズムの一般的な特徴
  - 尤度の単調性による安定した振る舞い
  - インプリメンテーションの容易さ
  - 適用範囲の広さ
    →不完全な観測として捉えることのできる
    統計モデルの推定に適用することができる
- EMアルゴリズムの欠点
  - 必ずしも大域的最適解に収束するとは限らない
  - 収束が遅い