

ベイズ統計入門 & モンテカルロ法と逆問題

伊庭幸人

I. ベイズ推定

条件付き確率と同時確率

x: 緑色の眼 **y**: 栗色の髪

$P(x|y)$ 髪が栗色のときに、眼が緑色の確率

$P(x, y)$ 髪が栗色で、かつ、眼が緑色の確率

$$P(x, y) = P(x|y)P(y)$$

$$P(x, y) = P(y|x)P(x)$$

ベイズの公式(ベイズの定理)

$$P(x|y)P(y) = P(x, y) = P(y|x)P(x)$$

$$P(y) = \sum_x P(x, y) = \sum_x P(y|x)P(x)$$

$$P(x|y) = \frac{P(x, y)}{\sum_x P(x, y)} = \frac{P(y|x)P(x)}{\sum_x P(y|x)P(x)}$$

例1



外から声をかけたら女性の声でした。
さてどちらの部屋だろうか。

順方向の確率

Aの部屋

女性の声 $\frac{1}{2}$ $P(F|A) = \frac{1}{2}$

男性の声 $\frac{1}{2}$ $P(M|A) = \frac{1}{2}$

Bの部屋

女性の声 1 $P(F|B) = 1$

男性の声 0 $P(M|B) = 0$

ベイズの公式を使う

$$P(x|y) = \frac{P(y|x)P(x)}{\sum_x P(y|x)P(x)}$$

女性の声

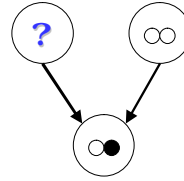
$$P(A|F) = \frac{P(F|A)P(A)}{P(F|A)P(A) + P(F|B)P(B)}$$

「事前確率」を $P(A) = P(B) = 1/2$ と仮定すると

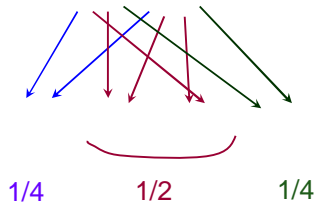
$$P(A|F) = \frac{(1/2)}{(1/2) + 1} = 1/3 \quad P(B|F) = 2/3$$

事後確率

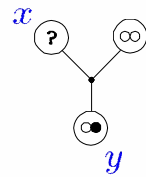
例2 メンデルの法則を逆に使う



メンデルの法則



メンデルの法則 条件付き確率



$$P(y = O● | x = OO) = 0$$

$$P(y = O● | x = O●) = \frac{1}{2}$$

$$P(y = O● | x = ●●) = 1$$

ハーディ・ワインバーグ平衡

●と○の頻度の比が $q : 1 - q$

$$P(x = OO) = (1 - q)^2$$

$$P(x = O●) = 2q(1 - q)$$

$$P(x = ●●) = q^2$$

両者を組み合わせる

●と○の頻度の比が $q:1-q$

$P(x = \text{○○}) = (1-q)^2$
 $P(x = \text{○●}) = 2q(1-q)$
 $P(x = \text{●●}) = q^2$

$P(y = \text{○●} | x = \text{○○}) = 0$
 $P(y = \text{○●} | x = \text{○●}) = \frac{1}{2}$
 $P(y = \text{○●} | x = \text{●●}) = 1$

ベイズの公式

$$P(x|y) = \frac{P(y|x)P(x)}{\sum_x P(y|x)P(x)}$$

$P(x = \text{●●} | y = \text{○●})$
 $= \frac{1 \times q^2}{\frac{1}{2} \times 2q(1-q) + 1 \times q^2} = q$

$q = 1/2$: ●●の確率は $\frac{1}{2}$
 ●がとても稀な遺伝子で $q = 10^{-4}$: 10^{-4}

実は事前分布を除くと同じ

Diagram 1: A node containing (●●) or (○●) has two children: (○○) and (○●). A red arrow points from the node to the child (○●).
 Diagram 2: A node containing (女性 男性) or (女性 女性) has two children: (●) and (○). A red arrow points from the node to the child (○).

例3. 最小2乗法

普通の「最小2乗法」をベイズで導出してみる

雑音の分布を設定

$y_i = y_i^* + \eta_i$ y_i^* 真の値
 η_i 独立&分散 σ^2 の正規分布に従う

$$P(y_i | y_i^*) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(y_i - y_i^*)^2}{2\sigma^2}\right)$$

条件つき確率の式

$y_i^* = at_i^* + b$ t_i は観測点の横軸の値(既知)
 $y_i = at_i^* + b + \eta_i$

$$P(\{y_i\} | a, b) = \prod_i \left\{ \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(y_i - at_i^* - b)^2}{2\sigma^2}\right) \right\}$$

$$P(\{y_i\} | a, b) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{1}{2\sigma^2} \sum_i (y_i - at_i^* - b)^2\right)$$

事前分布・事後分布

$P(a), P(b)$ 一様分布(一様密度とする)
 $P(\{y_i\} | a, b) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{1}{2\sigma^2} \sum_i (y_i - at_i^* - b)^2\right)$

ベイズの公式でひっくり返す

$$P(a, b | \{y_i\}) \propto \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{1}{2\sigma^2} \sum_i (y_i - at_i^* - b)^2\right)$$

最小2乗推定を再現

$$P(a, b | \{y_i\}) \propto \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{1}{2\sigma^2} \sum_i (y_i - at_i^* - b)^2\right)$$

最大にする a, b を求める

$$\sum_i (y_i - at_i^* - b)^2 \text{ を最小にする} \quad \text{最小2乗推定}$$

この場合事後分布での a, b の期待値を考えても同じ
(正規分布の特徴; 一般 \times)

つっこみどころ満載

$$P(a), P(b)$$

なんで a, b が一様分布で $a^3, \exp(b)$ はそうじゃないの？

そもそも無限区間なんだから
一様「分布」じゃないじゃん

事前分布の設定をめぐる困難

事前分布の設定は古来議論が絶えない

「ベイズ統計学」18世紀に成立

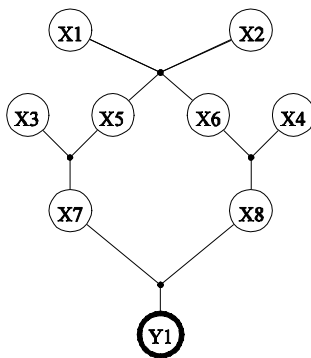
新理論というわけではない。また300年ぶりに再登場した
わけでもない(戦後に一度ブームあり)

統計学の「原型」的なもの

「アンチ事前分布」 推測統計学の成立

本講義はアルゴリズム中心の 最尤推定
講義なのでこれ以上突っ込まない 検定論

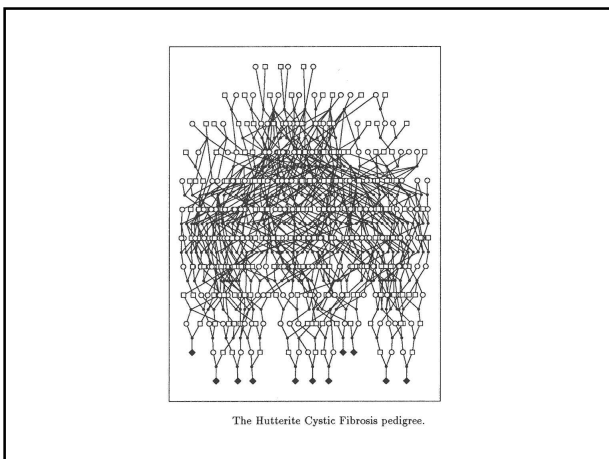
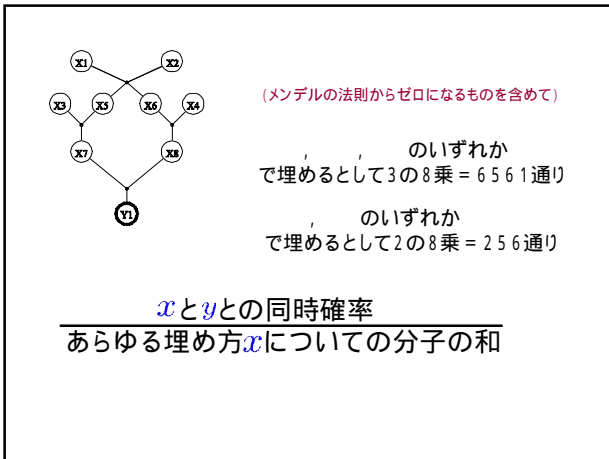
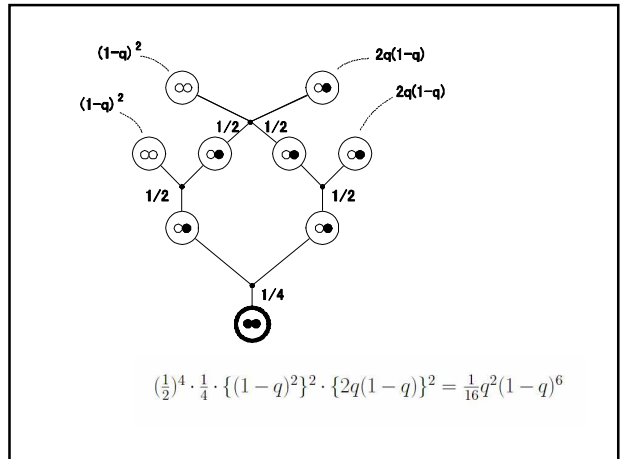
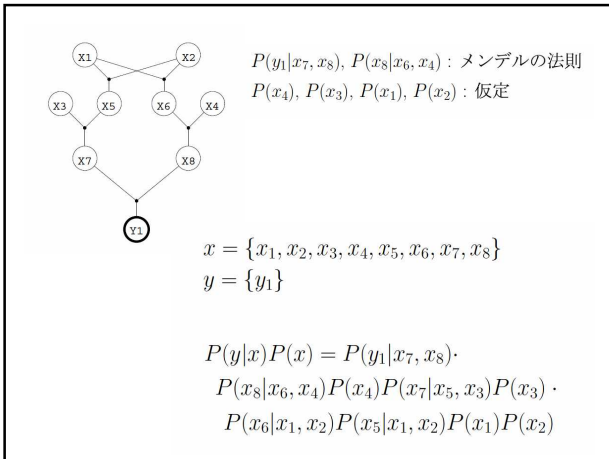
II. MCMCが必要なわけ



ベイズの公式

$$P(x|y) = \frac{P(x, y)}{\sum_x P(x, y)} = \frac{P(y|x)P(x)}{\sum_x P(y|x)P(x)}$$

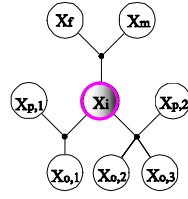
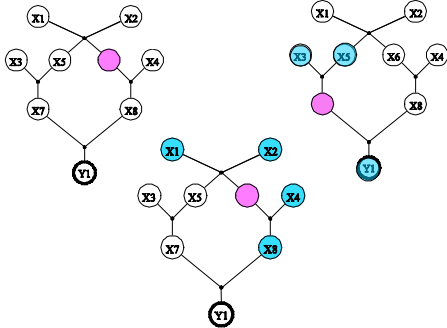
x と y との同時確率
あらゆる埋め方 x についての分子の和



実際にはたいへん
 埋める場所が100 3の100乗
 ベイズの公式は簡単だが、
 実際の計算はたいへん。

とくに、ループがあると簡単にできない
 (ループが少ないとき
 転送行列法に相当する手法で解ける)

ギブス・サンプラーを適用



$$P(x_i | \{x_j\}_{j \neq i}, y) = \frac{P(x_i | x_m, x_f) \prod_{o,p} P(x_i | x_p, x_o)}{\sum_{x_i} P(x_i | x_m, x_f) \prod_{o,p} P(x_i | x_p, x_o)}$$

遺伝子伝播だけでなく...

一般に
確率的な関係をネットワークでモデル化
ベイジアン・ネット(グラフィカル・モデル)

例: プリンタの故障診断

同様な「推論」はいろいろなところに現れる

一般化

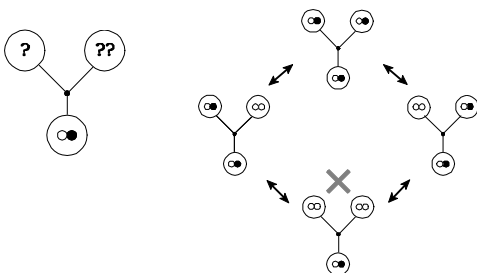
どのように込み入った, 要素数の多い,
推論のネットワークでも

「局所的な構造」

が簡単なら, 同様にして扱える!!

万能推論装置としてのモンテカルロ法

実はそんなに簡単ではない



「たくさんの系」を用意する

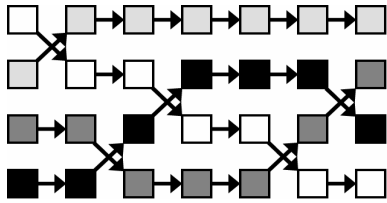
温度が少しずつ段階的に違うようにする

$$P_k(x^k) = \frac{\{P(x^k)\}^{1/T_k}}{\sum_x \{P(x^k)\}^{1/T_k}}$$

遺伝の場合は温度ではだめ 制約をゆるめる
任意のパラメータ についての族でもよい

$$P_k(x^k) = P(x^k | \beta_k)$$

交換の様子
(温度の異なる系4つの場合)



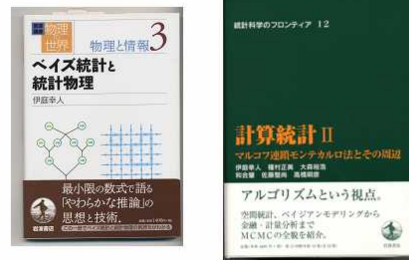
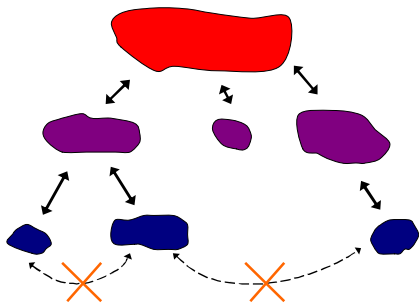
交換の規則

- 以下の比を計算

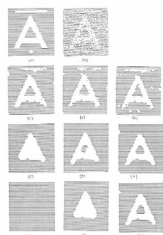
$$W = \frac{P_{k+1}(x^{k(t)}) P_k(x^{k+1(t)})}{P_{k+1}(x^{k+1(t)}) P_k(x^{k(t)})}$$

- 一様乱数 $0 < rnd < 1$ を発生
 $rnd < W$ なら交換する

「橋」をかける



III. マルコフ場による画像復元



Informative Prior

「よくわからないから・・・にする」 ignorant prior
impartial prior

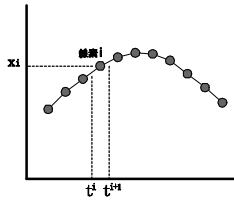
「事前分布に知識を入れてやる」 informative prior

平滑化事前分布 (~ 状態空間モデリング)

マルコフ場による画像復元

共役事前分布(~ 仮想データ): 大昔からある

平滑化

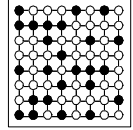


ガウス事前分布の範囲でも非常に有用
実用性ではイジング模型とは比較にならない

デモが楽しいのでイジングをやる

強磁性イジングモデル

- (i,j) は格子の上の隣接する対
- S_i と S_j が同符号ならエネルギー小
- S_i と S_j が異符号ならエネルギー大



$$P(S_1, S_2, \dots, S_N) = \frac{\exp(J \sum_{(ij)} S_i S_j)}{Z}$$

$$Z = \sum_{\text{config.}} \exp(J \sum_{(ij)} S_i S_j)$$

強磁性イジングモデルを事前分布にする

2値画像についての知識を表現する
「白の隣は白が多く、黒の隣は黒が多い」

$$P(S_1, S_2, \dots, S_N) = \frac{\exp(J \sum_{(ij)} S_i S_j)}{Z}$$

画素の条件付き確率

$$P(d_i | S_i)$$

$$d_i S_i = +1 \quad \text{with } 1-p$$

$$d_i S_i = -1 \quad \text{with } p$$

$$p(d_i | S_i) = \frac{\exp(h d_i S_i)}{\exp(h) + \exp(-h)}$$

$$h = \frac{1}{2} \log \frac{1-p}{p} \Leftrightarrow p = \frac{\exp(-h)}{\exp(h) + \exp(-h)}$$

データ全体の条件付き確率(尤度)

$$P(d_1, d_2, \dots, d_N | S_1, S_2, \dots, S_N)$$

$$= \prod_i \left\{ p^{\frac{1+d_i S_i}{2}} (1-p)^{\frac{1-d_i S_i}{2}} \right\}$$

$$= \frac{1}{Z'} \exp \left(h \sum_i d_i S_i \right)$$

$$h = \frac{1}{2} \log \frac{1-p}{p} \quad Z' = (\exp(h) + \exp(-h))^N$$

ベイズの公式を使う

$$P(S_1, S_2, \dots, S_N | d_1, d_2, \dots, d_N)$$

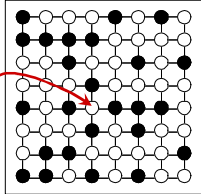
$$= \frac{1}{Z''} \exp \left(J \sum_{(ij)} S_i S_j + h \sum_i d_i S_i \right)$$

20x20 でも、まじめに扱うなら
2の400乗個の状態を相手にしなくてはならない

熱浴法(ギブス・サンプラー)

$$P(\cdot | S_1, S_2, \dots, S_{i-1}, S_{i+1}, \dots, S_N)$$

で S_i の値をランダムに取りなおす



これを選ぶ
とする

ギブスサンプラー

$$H_i = J \left(\sum_{j \in N(i)} S_j \right) + h_i$$

データ (± 1)

$$P(S_i = +1 | \cdot) = \frac{\exp(H_i)}{\exp(H_i) + \exp(-H_i)}$$

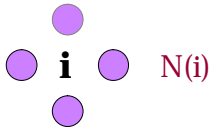
$$P(S_i = -1 | \cdot) = \frac{\exp(-H_i)}{\exp(H_i) + \exp(-H_i)}$$



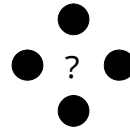
比較: ギブスサンプラー(データなし)

$$P(S_i = +1 | \cdot) = \frac{\exp(J \sum_{j \in N(i)} S_j)}{\exp(J \sum_{j \in N(i)} S_j) + \exp(-J \sum_{j \in N(i)} S_j)}$$

$$P(S_i = -1 | \cdot) = \frac{\exp(-J \sum_{j \in N(i)} S_j)}{\exp(J \sum_{j \in N(i)} S_j) + \exp(-J \sum_{j \in N(i)} S_j)}$$



条件付き確率の計算例



$$P(S_i = +1) = \frac{\exp(4J + hd_i)}{\exp(4J + hd_i) + \exp(-4J - hd_i)}$$

$$P(S_i = -1) = \frac{\exp(-4J - hd_i)}{\exp(4J + hd_i) + \exp(-4J - hd_i)}$$

情報の取り出し

「動いているのを見る」という方法もあるが

方法1. 画素ごとに独立に決める

S_i が1の回数が多ければ1, -1の回数が多ければ-1

MPM推定値

方法2. アニーリング

(事後) 確率の最大になる画像をもとめる
MAP推定値

ベイズ最適解 $\{\tilde{S}_i\}$

$$f(\{\tilde{S}_i\}) = \sum_{\{s_i\}} \left\{ \sum_i \tilde{S}_i s_i \right\} P(\{S_i\} | \{d_i\})$$

$$= \sum_i \left\{ \tilde{S}_i \times \sum_{\{s_i\}} s_i P(\{S_i\} | \{d_i\}) \right\}$$

$$\tilde{S}_i = \text{sign} \left\{ \sum_{\{S_i\}} s_i P(\{S_i\} | \{d_i\}) \right\}$$

アニーリング

$$\beta \rightarrow \infty$$

$$\frac{P(\{S_i\}|\{d_i\})^\beta}{\sum_{\{S_i\}} P(\{S_i\}|\{d_i\})^\beta}$$

今の場合..

h, J をきめる方法がなければ
あんまり意味がない

事後分布が「ほんとう」

本当に事前分布で「生成」されたパターン
に大きさ/性質が既知の雑音を加えたとき

イジングのデモ

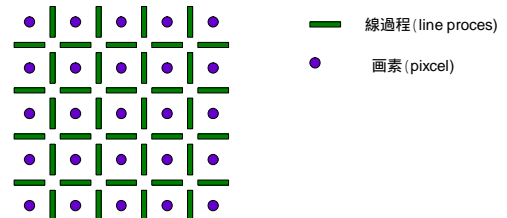
[javatest¥Bayes.html](#)

[javatest¥Bayes4.html](#)

[javatest¥ABayes.html](#)

[javatest¥Bayes00.html](#)

線過程を含むマルコフ場事前分布



画素の相互作用を線過程が制御 境界を定義する
線過程同士も相互作用 境界がまっすぐで枝分かれが少ない

線過程の「エネルギー」

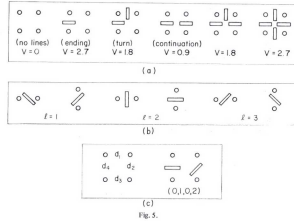


Fig. 3.

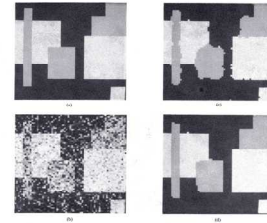


Fig. 4. (a) Original image. (b) Hard-Down. (c) Soften image. (d) Soften image. Without the original and position. (e) Estimated image. Approximate. (f) Estimated image.

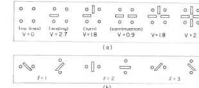


Fig. 5.

20年前に評判になった理由

ベイズ + マルコフ場 + ギブスサンプラー + SA
わかりやすい成果 脳/認知研究とも関連付け

協力現象, 自己組織化といったキーワード
と実用的な処理をつなげてみせた

プログラム可能性: 知識を容易に作りこめる

問題点

$\beta h, \beta J$ などの(ハイパー)パラメータを
教師なしで学習する筋道だった手法が未確立
「事前分布」が単独で意味を持たない点に関連. ベイズの
一番ディープな闇に触れる?

応用がそれほどない(?)

大域的情報 例: 顔認識

運動に重点 例: トラッキング

その後の展開

工学系の中ですでにマニアックになって
狭い「業界」を作ってしまったようなところがある

物理学者 さまざまな近似法を導入

それらはまったく本質的な問題ではない

問題点が別にあるのに

計算手法を工夫してもだめ

今年の子定

東工大での講義(後期, たぶん2~3回分)

20日の分+ +

計算科学夏の学校

9月27日~29日(2日目の午前)

電子情報通信学会 NC研究会(奈良)

10月 パネル討論?

総研大での授業

後期の開講は未定

統致研
(総研大統計科学専攻)
学生募集中
博士3年間 & 5年一貫
学振ももちろん出せませす

MCMCの統計への応用

代表的なヒットは何? 難しい

- @ 統計・情報の中でも
それぞれの分野で別々にやっている
- @ この講義では触れないが, 少し違う
「逐次モンテカルロ」も流行

最後のあまけに簡単なのをやる

マーケティング・経済時系列・疫学・生態学

系統樹の解析

トラッキング, ナビゲーション (逐次モンテカルロ)

以下では..

物理学科でもあるし, 物理と情報?
の境界ということで, 自分の興味で
やや趣味的に走る

自分の研究の紹介も入れる

V. 逆問題とベイズの公式いろいろ

「逆に解く」ことの重要性

「ベイズの公式」によれば
一般に「データの生成過程」のモデル
があればそれを反転して解くことができる

- 例 先祖の持つ遺伝子 子孫の遺伝子
- 例 事故の原因 事故の内容
- 例 顧客の心理 購買行動 (?)

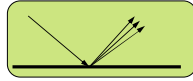
アニメーションの自動生成への応用

S Cheney

珍しい現象 たとえば「ホールインワン」とかの
絵を自動的に描く

「手でかけばいいじゃん」という突っ込みはなし

単なるシミュレーション



例: $P(Y | X)$

X: 地面に当たる角度 Y: 跳ね返る角度
Y: 鏡映対称から ± 10 度で一様分布

$$P(D | C)P(C | B)P(B | A)P(A)$$

A B C D と順次求めていけばよい

乱数はいるが、マルコフ連鎖モンテカルロ法はいらない

「珍しい結果」を出すには？

Chenney and Forsyth (2000)

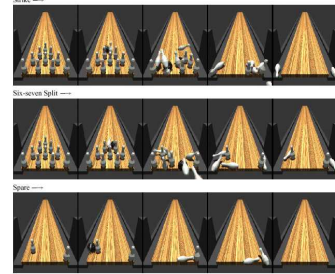


Figure 3. Frames from three bowling examples. The initial conditions for the ball and the pin locations are random variables. Given an initial and final pin configuration, the MCMC algorithm samples particular values for the random variables that lead to the desired shot. In this case, we demarcated a strike, a six-seven split and the corresponding spare.

逆に考える

$$\begin{aligned} & P(D | C)P(C | B)P(B | A)P(A) \\ = & P(A, B, C, D) \\ = & P(A, B, C | D)P(D) \end{aligned}$$

ベイズの公式: Dを固定したときの条件つき確率

$$\begin{aligned} & P(A, B, C | D) \\ = & P(D | C)P(C | B)P(B | A)P(A) / \underbrace{P(D)}_{\text{分母は定数}} \end{aligned}$$

珍しい現象のアニメを作る

重み $P(D|C)P(C|B)P(B|A)P(A)$ で
「Dが実現されるまでの事象の系列A, B, C」
をサンプリングすればよい

マルコフ連鎖モンテカルロ法を利用

系列全体を変数として少しずつ変形しながらシミュレーション

メトロポリス法

対称な提案分布 $Q(x, x') = Q(x', x)$

x の候補 x^{new} を
提案分布 $Q(x, x^{new})$ に従って生成

一様乱数 $0 \leq r < 1$ を発生

$r < \frac{P(x^{new})}{P(x)}$ なら x を x^{new} で置き換える
さもなければ何もしない

メトロポリス・ヘイスティングス法 (MH)

$$Q(x, x') \neq Q(x', x)$$

$$r < \frac{P(x^{new})Q(x^{new}, x)}{P(x)Q(x, x^{new})}$$

なら x を x^{new} で置き換える
さもなければ何もしない

「珍しい結果」を出すには？

Chenney and Forsyth (2000)

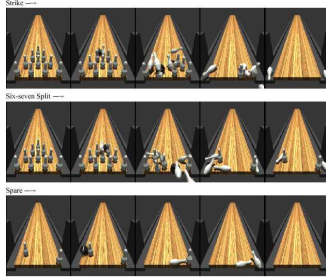


Figure 3. Frames from three bowling examples. The initial conditions for the ball and the pin locations are random variables. Given an initial and final pin configuration, the MCMC algorithm samples particular values for the random variables that lead to the observed shot. In this case, we demanded a strike, a six-seven split and the corresponding spare.



Chenney and Forsyth (2000)

珍しい現象のサンプリング

ふたたび物理の世界へ...

普通にシミュレーションしていたのでは滅多に起きない現象
分子の形態がAからBに変化

$$P(B | Z)P(Z | Y)P(Y | X)P(X | A)P(A)$$

両端A, Bを固定してベイズの公式で

途中の状態Z, Y, Xを求める

以下 Bolhuis et al. (2000) より

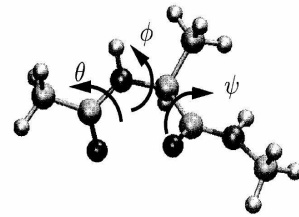


Fig. 1. Schematic representation of the alanine dipeptide molecule ($\text{CH}_3\text{-CONH-CH(CH}_3\text{-CONH-CH}_3$). The backbone dihedral angles are labeled ϕ and ψ . The additional variable θ is the torsional angle O-C-N-C_α , and is not to be confused with the dihedral angle α $\text{C}_\alpha\text{-C-N-C}_\alpha$.

58

DAVID CHANDLER

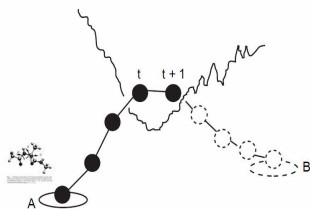
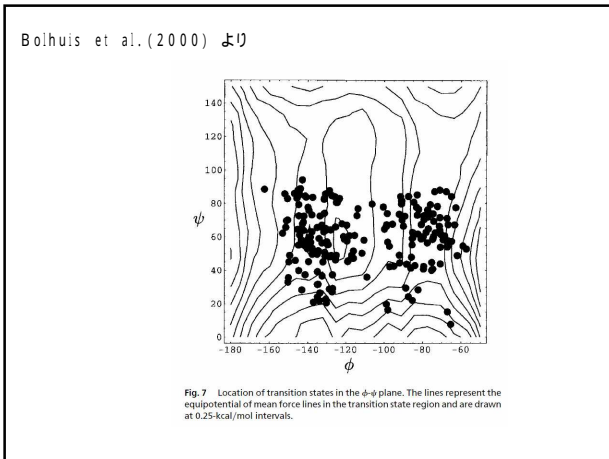
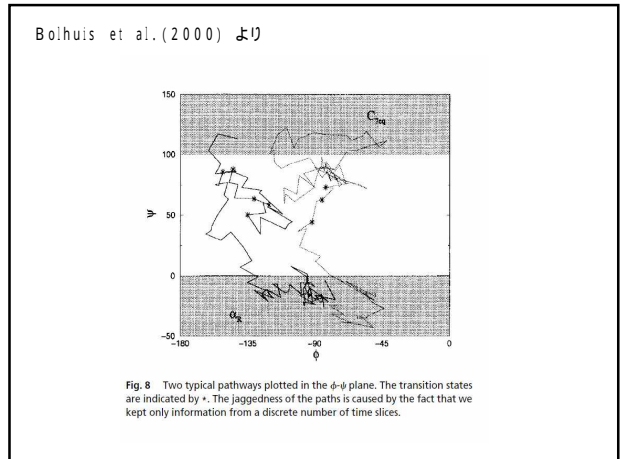
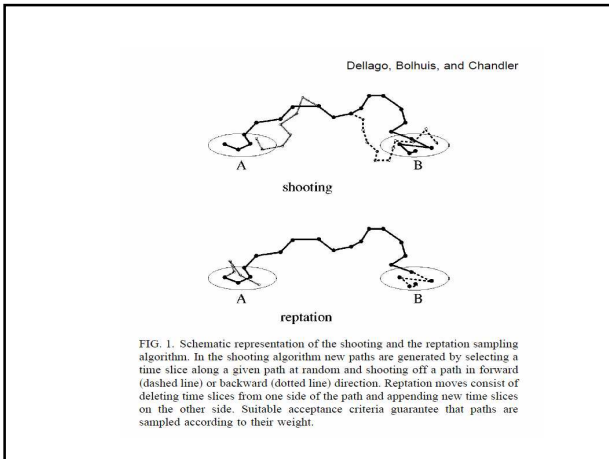


Fig. 4. Schematic pathway between stable states A and B, with points on the path marked at discrete points in time.

Bolhuis et al. (2000) より

$$\begin{aligned} & \exp(-S_{AB}[\{x\}]) \\ & \equiv h_A(x_0) e^{-\beta H(x_0)} \left[\prod_{\tau=0}^{L-1} p(x_\tau \rightarrow x_{\tau+1}) \right] h_B(x_L). \end{aligned} \quad (2)$$

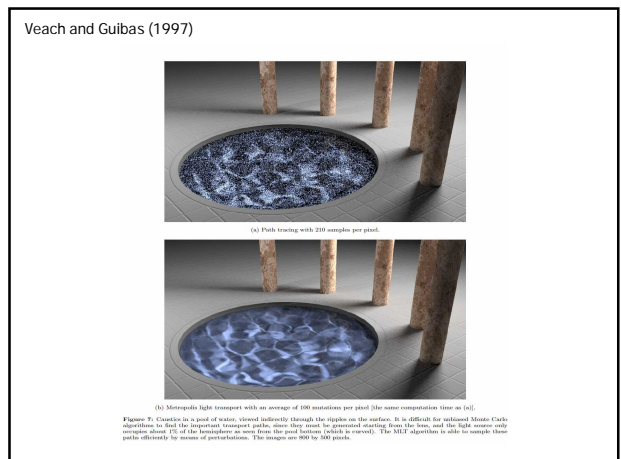
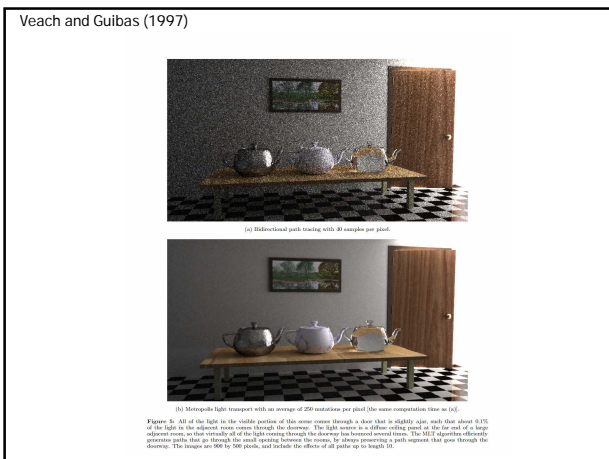


コンピュータグラフィクスへの応用

レンダリング (光線の追跡) Eric Veach

Photorealistic Rendering

表面形状, 反射関数と光源を与えて,
超リアルな画像を作りだす



なぜ難しい？

多重反射を正確に表現するには
シミュレーションが必須

光源から順方向にシミュレーションすると
ほとんどの「光線」は目に入らず無駄になる

目から逆に追った光線 ・ ・ 光源から出る光線

出会う確率はとても小さい

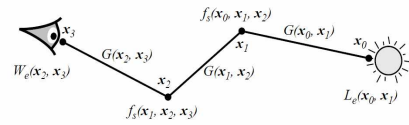


Figure 1: The measurement contribution function f_j is a product of many factors (shown for a path of length 3).

$$m_j = \int_{\mathcal{M}^2} L_e(\mathbf{x}_0 \rightarrow \mathbf{x}_1) G(\mathbf{x}_0 \leftrightarrow \mathbf{x}_1) W_e^{(j)}(\mathbf{x}_0 \rightarrow \mathbf{x}_1) dA(\mathbf{x}_0) dA(\mathbf{x}_1) \\ + \int_{\mathcal{M}^3} L_e(\mathbf{x}_0 \rightarrow \mathbf{x}_1) G(\mathbf{x}_0 \leftrightarrow \mathbf{x}_1) f_s(\mathbf{x}_0 \leftrightarrow \mathbf{x}_1 \leftrightarrow \mathbf{x}_2) \\ G(\mathbf{x}_1 \leftrightarrow \mathbf{x}_2) W_e^{(j)}(\mathbf{x}_1 \rightarrow \mathbf{x}_2) dA(\mathbf{x}_0) dA(\mathbf{x}_1) dA(\mathbf{x}_2) \\ + \dots \quad (8)$$

メトロポリス光輸送 (MLT)

両端固定の「経路」(光線の通り道)の上の測度を
マルコフ連鎖モンテカルロ法でサンプリング

経路を変形 サイコロを振って受理・棄却

物理で量子力学(場の理論)の
経路積分の計算(の確率論版)に使われている方法
(Stochastic Series Expansion)と同じ。

経路の変形

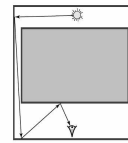


Figure 2: If only additions and deletions of a single vertex are allowed, then paths cannot migrate from one side of the barrier to the other.

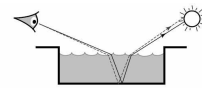
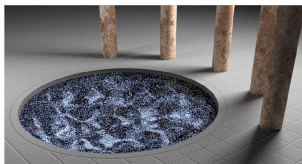
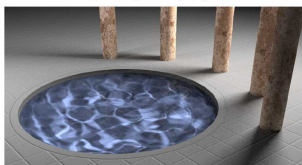


Figure 4: Using a two-chain perturbation to sample caustics in a pool of water. First, the lens edge is perturbed to generate a point x' on the pool bottom. Then, the direction from original point x toward the light source is perturbed, and a ray is cast from x' in this direction.

Veach and Guibas (1997)



(a) Path tracing with 250 samples per pixel.



(b) Metropolis light transport with an average of 100 mutations per pixel (for same computation time as (a)).

Figure 7: Caustics in a pool of water, viewed indirectly through the poles on the surface. It is difficult for unbiased Monte Carlo simulation to fill in the bright caustic region, since light must be generated starting from the lens and the light source casts rays only into the lens. The caustic region is rendered by the MLT algorithm. The images are 900 by 900 pixels.

VI.物理学会2006秋

* プレビュー *

秋の物理学会

25aXF-3 拡張アンサンブルモンテカルロによる力学系の特別な初期値探索
柳田達雄、伊庭幸人

25aXF-4 拡張アンサンブル法によるカオスサドルのサンプリング
伊庭幸人 柳田達雄

23aXD-11 グリフィス相の直接数値計算
福島孝治、伊庭幸人

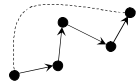
力学系の研究へのMCMCの応用

CGじゃなくて、
力学系の研究に応用できないか。

- @ 不安定周期軌道
- @ リペラーとかサドルとか
- @ カオスの統計力学
- @ 各種の大偏差統計

モンテカルロを用いた既存の研究

@ 「軌道」を状態変数としてMCMC



Doll ほか (1994)
Sasa, Kawasaki ほか (2005 ~)
Chandler Group (1998年ごろ ~)
Iba 集中講義 (駒場, 2003) など

@ もとのシステムの状態変数

逐次モンテカルロ Kurchanたち (2005)
(GAのように粒子が増減)

初期状態アンサンブル

deterministic な系専用 25aXF-3, 4

- @ 軌道を初期値で表現
- @ 求める「珍しい性質」「大偏差」に応じてエネルギー E を定義
- @ ギブス分布を考える
(「 $E < 0$ の一様分布」などでもよい)

以下の例では **escape time** でエネルギーを決める。

ジュリア集合のサンプリング

ある(範囲)の初期値から出発して軌道が無限遠にいかない初期値 z_0 の集合の境界

$$z = z^2 + C$$

「長ーく迷ってから無限遠に去る z_0 の集合」
をMCMCで探求する

ジュリア集合のMCMC

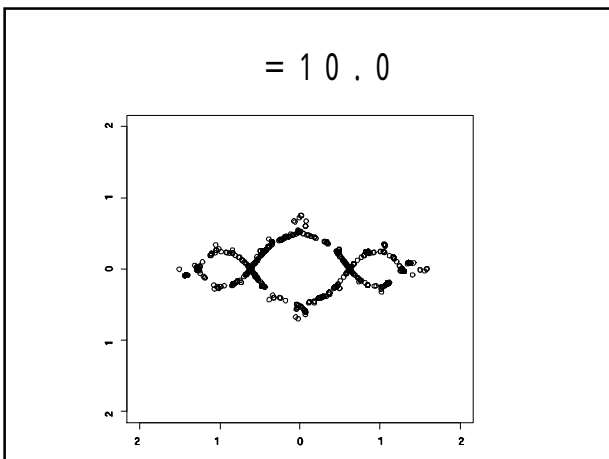
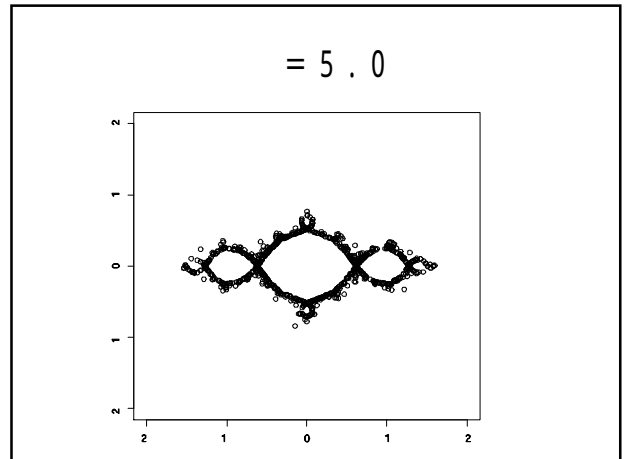
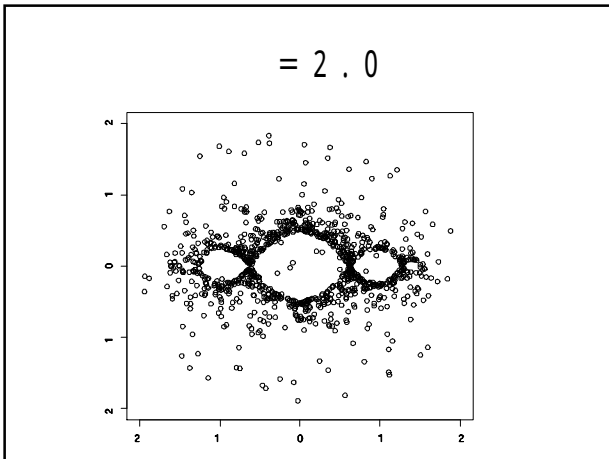
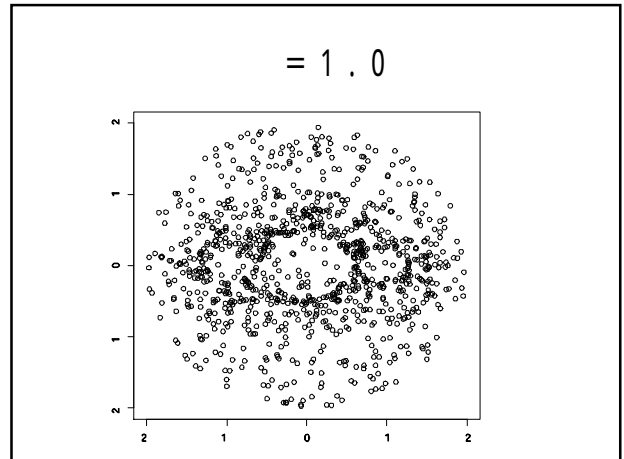
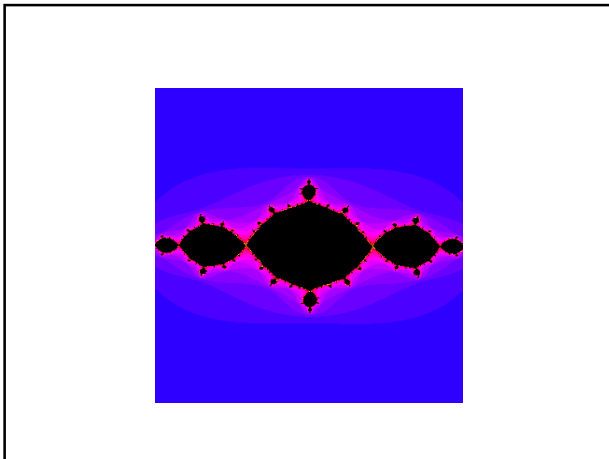
「抜けるまでのステップ数」 n

$$U(z_0) = \log n \text{ (if } n > c : U = 0)$$

$$P(z_0) = \exp(-U(z_0))/Z$$

メトロポリス法 + パラレルテンパリング

もっと初期値が高次元の場合には実用性？

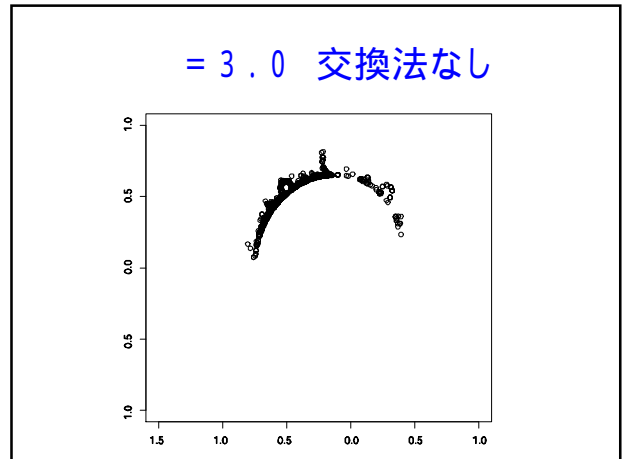
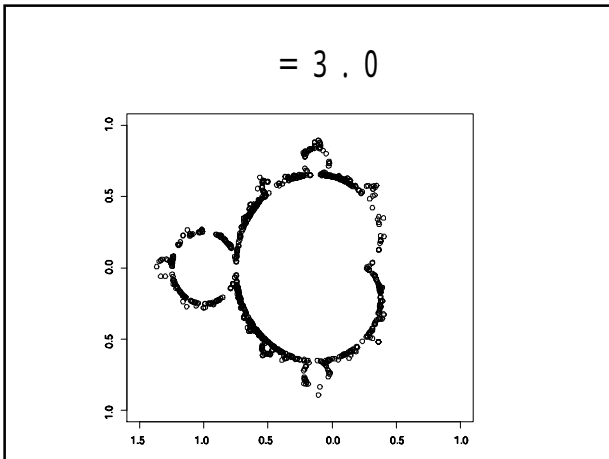
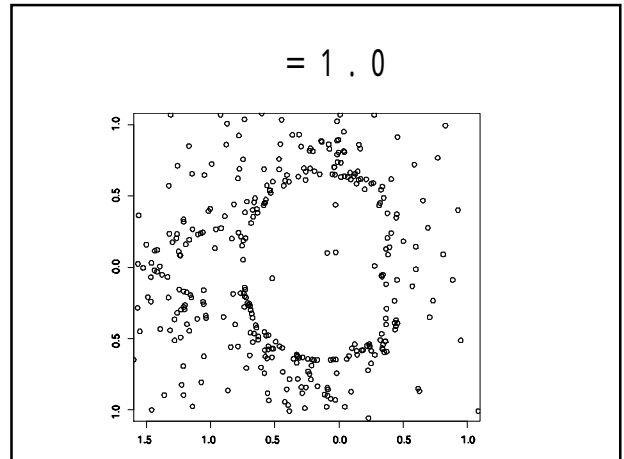
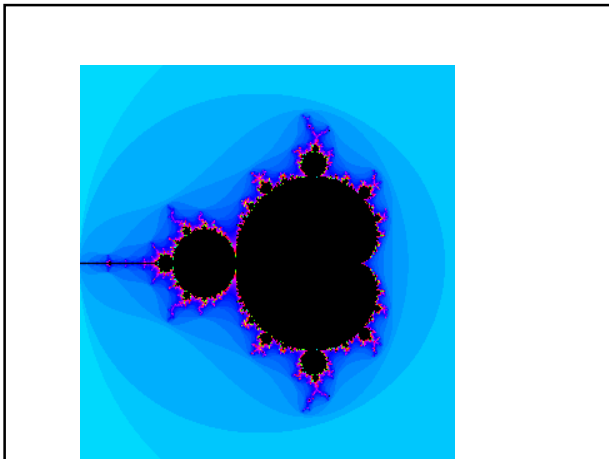


マンデルブロ集合の縁

ある(範囲)の初期値から出発して軌道が無限遠にいかないパラメータCの集合

$$z = z^2 + C$$

「長ーく迷ってから無限遠に去るCの集合」をMCMCで探求することもできる



ここまで
物性研究85 3 (2005年12月号)
in 「モンテカルロ法の新展開3」報告

このあと
柳田達雄氏(北大電子研)との
共同研究

Stagger and Step Method

D. Sweet, H. E. Nusse, and J. A. Yorke
Phys Rev Lett 86 2261-2264 (2001) (末谷さんが教えてくれた)

高次元カオスでのサドルを「見る」手法

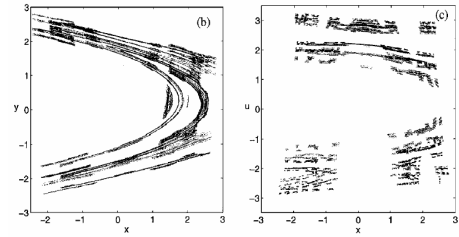
- 1) 良い軌道をescape timeを最大化することで求める
探索時の動かし方にコツあり : 階層的なmove
- 2) 飛び出したらランダムに微小量ずらしてやりなおす.
「待たあり」のシミュレーション 擬軌道

結合エノン写像 4次元のリペラー

$$\begin{aligned}
 x_{t+1} &= a - x_t^2 + by_t + k(x_t - u_t) \\
 y_{t+1} &= x_t \\
 u_{t+1} &= c - u_t^2 + dv_t + k(u_t - x_t) \\
 v_{t+1} &= u_t
 \end{aligned}$$

$a = 3, c = 5, b = d = 0.3, k = 0.4$

彼らの結果



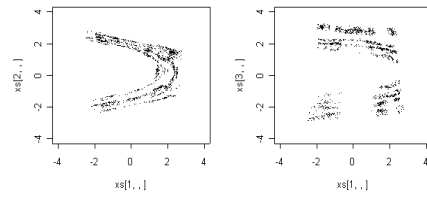
Stagger and Step法による擬軌道

初期値モンテカルロ

- @ ギブス分布
E = -log(escape time)
- @ 閾値アンサンブル
P (初期値) =
[escape time >] での一様分布

メトロポリス + パラレルテンパリング

モンテカルロで求めたリペラー
(結合エノン 4次元)

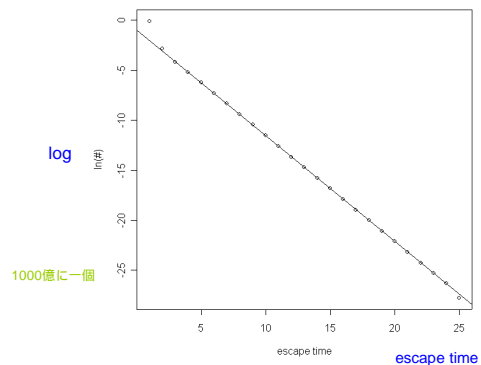


初期値アンサンブルから10回程度iterate

比較

- @ 4次元でも一応できた
- @ 一個のリペラーを「見る」だけなら Stagger and Stepでいいかも ..
- @ 統計量でも出してみるか..

初期値集合上のescape timeの分布
リペラー (結合エノン 4次元)



suetani's special

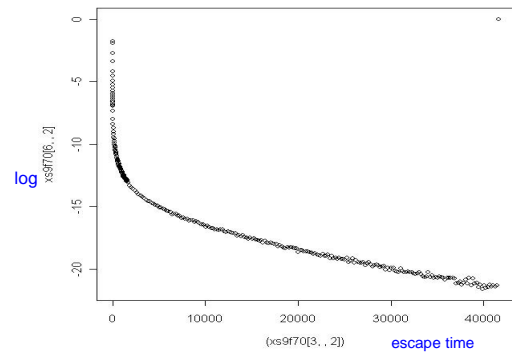
$$x_{t+1} = 4x_t(1 - x_t)$$

$$y_{t+1} = \nu e^{-5(x_t - 0.75)^2} y_t + x_t^3$$

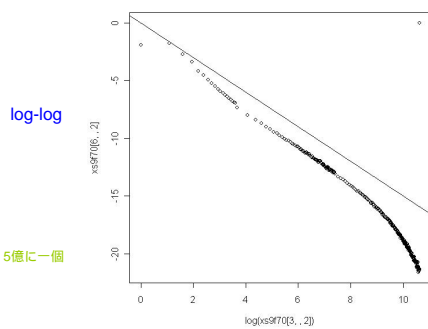
riddled basin $y = 0$ or $y = 0$: < 2.553522
 $y = 0$ が不安定化 : > 2.553522

指数3/2のベキ 指数関数 cross over

エスケープレート riddled basin 崩れ =2.6



エスケープレート riddled basin 崩れ =2.6



- @ なんか地味かも～
- @ 面白い例はないか？

柳田さんと話す

「2重振り子の逆立ちはどうか」

2重振り子 by 柳田

[pen23.avi](#)

まとめ

「多体問題」的な確率的問題は物理以外のいたるところにみられる

動的モンテカルロ法などの統計物理の手法が多くの分野で注目されている

「逆に解く」「珍しい現象を積極的に作る」という思想は物理学においても有効

補論. 欠測の補完

(狭義の)ベイズ統計 + MCMCの典型例

複雑な打ち切り (truncation) の例

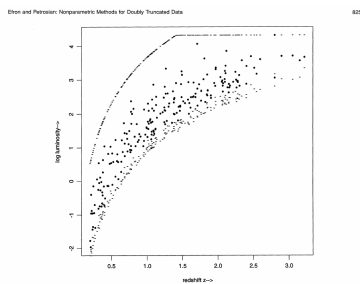
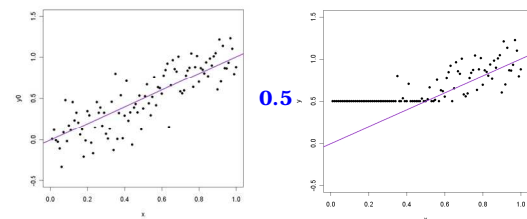


Figure 1. Cloudy Truncated Data. Points represent redshifts and luminosities for 210 quasars, as described in Section 6. Luminosity subject to lower and upper truncation as indicated by dashed lines (lower truncation fixed at 25 for clarity) is uniformly correlated with redshift?

Efron and Petrosian (JASA)

ここで考える例



ここでの方針

直観から出発する MCMCまがいの手法

ベイズの枠組みで定式化

最初のものに似た(完全に同じではないが)手法が系統的に導けることを示す

素朴にやると...

その1

打ち切りデータの y を打ち切りの上限(0.5)としてあてはめる

その2

打ち切られていないデータだけを使う

どちらも回帰直線が寝てしまう

デモをやる.

では, どうするか.. (方法その3)

まず直線をあてはめる $y = ax + b$
残差から雑音の分散 σ^2 を求める

打ち切り部分のデータをシミュレートする
 $y_i = ax_i + b + (\text{分散 } \sigma^2 \text{ のガウス乱数})$

それに直線をあてはめ, 分散を求める
以下これを繰り返す

「見えている」データはずっとそのまま

デモをやる.

なぜうまく行かないか

値が ~ 0.5 だとする

+ 雑音で

0.5より大 見えている 50%

0.5より小 打ち切られる 50%

打ち切られた部分をシミュレート

0.5より大, 小 25%ずつ

合計75%が0.5より大になる バイアス

では, どうするか.. (方法その4)

まず直線をあてはめる $y = ax + b$
残差から雑音の分散 σ^2 を求める

打ち切り部分のデータをシミュレートする

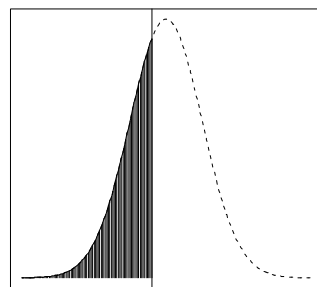
$y_i = ax_i + b +$

は分散 σ^2 の切断正規分布に従う

$0.5 > ax_i + b +$ $0.5 - ax_i - b >$

あとの繰り返しは同じ

切断正規分布(truncated normal)



デモをやる.

これで一応よさげだが..

全体として何をやっているのか
途中から暴走したりしないのか

いつまでも動いていることの意味

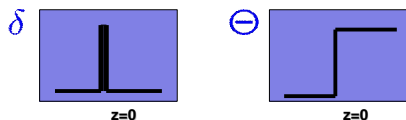
a, bや σ^2 の誤差が見積もれないか

ベイズ統計 + Gibbs Samplerにもとづく解釈

階層ベイズ風の定式化(1)

$$p(\{y_i\}|\{z_i\}) = \prod_{i \in A} \delta(y_i - x_i) \\ = \prod_{i \notin A} \Theta(0.5 - z_i)$$

z_i hidden variable; $A = \{z_i | y_i > 0.5\}$



階層ベイズ風の定式化(2)

$$p(z_i|a, b, \sigma^2) = \left(\frac{1}{2\pi\sigma^2}\right)^{N/2} \exp\left\{-\frac{(z_i - ax_i - b)^2}{2\sigma^2}\right\}$$

$p(a), p(b), p(\sigma^2)$ とりあえず定義域で一様密度

本当はおかしい(規格化不能)

事後分布

$$p(\{y_i\}|\{z_i\}, a, b, \sigma^2) \\ = p(\{y_i\}|\{z_i\})p(\{z_i\}|a, b, \sigma^2)p(a)p(b)p(\sigma^2)$$

$$p(\{z_i\}, a, b, \sigma^2|\{y_i\}) \\ \propto p(\{y_i\}|\{z_i\})p(\{z_i\}|a, b, \sigma^2)p(a)p(b)p(\sigma^2)$$

$$= \left(\prod_{i \in A} p(y_i|a, b, \sigma^2)\right) \left(\prod_{i \notin A} \Theta(0.5 - z_i)p(z_i|a, b, \sigma^2)\right) p(a)p(b)p(\sigma^2)$$

ギブスサンプラー(1)

$i \notin A$

$$z_i \leftarrow p(z_i|a, b, \sigma^2)$$

$$z_i \leftarrow \frac{\Theta(0.5 - z_i)p(z_i|a, b, \sigma^2)}{\int \Theta(0.5 - z'_i)p(z'_i|a, b, \sigma^2) dz'_i}$$

切断正規分布

$$p(z_i|a, b, \sigma^2) = \left(\frac{1}{2\pi\sigma^2}\right)^{N/2} \exp\left\{-\frac{(z_i - ax_i - b)^2}{2\sigma^2}\right\}$$

ギブスサンプラー (2)

$$p(a|\{z_i\}, b, \sigma^2) \propto \exp\left\{-\frac{\sum_{i \in A} (y_i - ax_i - b)^2}{2\sigma^2}\right\} \exp\left\{-\frac{\sum_{i \notin A} (z_i - ax_i - b)^2}{2\sigma^2}\right\}$$

$$p(a|\{y_i\}, \{z_i\}, b, \sigma^2) \propto \exp\left(-\frac{1}{2\sigma_*^2}(a - \hat{a})^2\right)$$

\hat{a} は $\{y_i\}, \{z_i\}$ を与えたときの a の最小二乗推定値
 $\sigma_*^2 = \sigma^2 / \sum_i x_i^2$

ギブスサンプラー (3)

$$p(b|\{z_i\}, a, \sigma^2) \propto \exp\left\{-\frac{\sum_{i \in A} (y_i - ax_i - b)^2}{2\sigma^2}\right\} \exp\left\{-\frac{\sum_{i \notin A} (z_i - ax_i - b)^2}{2\sigma^2}\right\}$$

$$p(b|\{y_i\}, \{z_i\}, a, \sigma^2) \propto \exp\left(-\frac{1}{2\sigma_*^2}(b - \hat{b})^2\right)$$

\hat{b} は $\{y_i\}, \{z_i\}$ を与えたときの b の最小二乗推定値
 $\sigma_*^2 = \sigma^2 / \sum_i x_i^2$

ギブスサンプラー (4)

$$p(\sigma^2|\{z_i\}, a, b) \propto \left(\frac{1}{\sigma^2}\right)^{N/2} \exp\left\{-\frac{\sum_{i \in A} (y_i - ax_i - b)^2}{2\sigma^2}\right\} \exp\left\{-\frac{\sum_{i \notin A} (z_i - ax_i - b)^2}{2\sigma^2}\right\}$$

$$p(\sigma^2|\{z_i\}, a, b) \propto \left(\frac{1}{\sigma^2}\right)^{\alpha+1} \exp\left(-\frac{\beta}{\sigma^2}\right)$$

$$\alpha = \frac{N}{2} - 1$$

$$\beta = \frac{1}{2} \sum_i (z_i - ax_i - b)^2$$

逆ガンマ分布

直観的な方法との比較

観測されないサンプル z_i を選ぶ部分 (1) はまったく同じ

a, b, σ^2 を再推定する部分 (2) (3) (4) では点推定 (最尤推定) の代わりに, サンプリングしているところが違う **誤差**

ギブス・サンプラー

どういう分布を定常にするか既知!

EMアルゴリズム

$$p(\{z_i\}, a, b, \sigma^2|\{y_i\}) \propto p(\{y_i\}|\{z_i\})p(\{z_i\}|a, b, \sigma^2)p(a)p(b)p(\sigma^2) \int p(\{y_i\}|\{z_i\})p(\{z_i\}|a, b, \sigma^2) dz_1 \cdots dz_N$$

a, b, σ^2 は点推定になるがわれわれの直観的方法とはちょっと違う

デモをやる.

事前分布について

$p(a), p(b), p(\sigma^2)$
 $p(a), p(b)$ 適当な平均・分散の正規分布
 $p(\sigma^2)$ 適当な β の逆ガンマ分布

共役事前分布

$$p(\sigma^2|\{z_i\}, a, b) \propto \left(\frac{1}{\sigma^2}\right)^{\alpha+1} \exp\left(-\frac{\beta}{\sigma^2}\right)$$
$$p(\sigma^2) \propto \left(\frac{1}{\sigma^2}\right)^{\alpha'+1} \exp\left(-\frac{\beta'}{\sigma^2}\right)$$

仮想データ
(あるいはデータベースの内容)
と解釈できる

$$p(\sigma^2|\{z_i\}, a, b)p(\sigma^2) \propto \left(\frac{1}{\sigma^2}\right)^{\alpha+\alpha'+2} \exp\left(-\frac{\beta'+\beta}{\sigma^2}\right)$$

正規分布同士でも同様に形が保たれる