ゼロからできるMCMC 正誤表

ページは全て第 1 刷のもの。

2 刷での修正、変更

ページ等	修正前	修正後
p.47, 問題 3.1 の解答	$k-1$ 回目までに	k 回目までに
89 ページ、問題 5-1 の解答	n ステップをまとめたあとでは詳細釣り合いが壊れてしまいますが	問題 4-2 の時と同じ意味で詳細釣り合いは成り立っていませんが
p.134, 問題 6.3 の解答	$x(\tau)$ の時間発展 $x(\tau+\Delta\tau)+\Delta\tau\cdot p\left(\tau+\frac{\Delta\tau}{2}\right)$	$x(\tau)$ の時間発展 $x(\tau)+\Delta\tau\cdot p\left(\tau+\frac{\Delta\tau}{2}\right)$
p.121	このやり方でマルコフ連鎖モンテカルロ法の条件が満たされていることを確認しましょう. マルコフ連鎖であることと既約性はほとんど自明でしょう. 非周期性は「y を固定して x を更新」と「x を固定して y を更新」をまとめて 1 ステップと思えば成立しています. 最後に詳細釣り合い条件を慎重に調べましょう.	「y を固定して x を更新」と「x を固定して y を更新」をまとめて 1 ステップと思うことにしましょう. このやり方でマルコフ連鎖モンテカルロ法の条件が満たされていることを確認します. マルコフ連鎖であることと既約性、非周期性はほとんど自明でしょう. 以下, 詳細釣り合い条件を慎重に調べます.
p.140	これを μ と σ の関数と考えて「$\{x_i\}$ が実現されるもっともらしさ」と解釈するのが尤度なのでした.	これを μ と σ の関数と考えて「$\{x_i\}$ がパラメーター μ と σ から実現されるという仮定のもっともらしさ」と解釈するのが尤度なのでした.
p.149–p.150, Eq. (7.29)	$P(p\|n,k)\cdot P(p\|n',k')$ $=p^k(1-p)^{n-k}\cdot p^{k'}(1-p)^{n'-k'}$ $=p^{k+k'}(1-p)^{(n+n')-(k+k')}$ $=P(p\|n+n',k+k')$	$P(p\|n,k)\cdot P(p\|n',k')$ $\propto p^k(1-p)^{n-k}\cdot p^{k'}(1-p)^{n'-k'}$ $=p^{k+k'}(1-p)^{(n+n')-(k+k')}$ $\propto P(p\|n+n',k+k')$
p.150	$P(p\|k)=p^{35}(1-p)^{35}$	$P(p\|k)\propto p^{35}(1-p)^{35}$
p.153	N_p は確率 p のコインの総数です.	N_p は確率 p のコインが選ばれた回数です.
p.153, Eq.(7.34)	$P(k)=\lim_{N\to\infty}\frac{\sum_p n_{k,p}}{N}=\frac{N_k}{N}$	$P(k)=\lim_{N\to\infty}\frac{\sum_p n_{k,p}}{N}$ $=\lim_{N\to\infty}\frac{N_k}{N}$

ページ等	修正前	修正後
p.153, Eq.(7.34)	$P(p) = \lim_{N\to\infty} \frac{\sum_k n_{k,p}}{N} = \frac{N_p}{N}$	$P(p) = \lim_{N\to\infty} \frac{\sum_k n_{k,p}}{N} = \lim_{N\to\infty} \frac{N_p}{N}$
p.160	$E_\pm$ を計算するには点 i と隣接する点のスピンとの相互作用だけ考えれば良いので	E_+ と E_- の差を計算するには点 i と隣接する点のスピンとの相互作用だけ考えれば良いので
p.173	移動時間の合計が決まります.	移動距離の合計が決まります.
p.181, Fig. 7.18	真ん中	中央
p.206	**nskip** はメトロポリス法のサンプル採取頻度です.	**nskip** はサンプル採取頻度です.
p.201	`https://github.com/masanorihanada/MCMC-Sample-Codes` からダウンロードできます. ライブラリ等は使用していないので, 通常の C あるいは C++のコンパイラーだけでコンパイル可能です.	C, C++で書かれたサンプルコードを `https://github.com/masanorihanada/MCMC-Sample-Codes` からダウンロードできます. ライブラリ等は使用していないので, 通常の C あるいは C++のコンパイラーだけでコンパイル可能です. Python3 のコードも同じ GitHub アカウントで提供します.
Appendix A, コード名のところで 4 箇所	Bayse	Bayes

p.75, 問題 4.2 の解答

まず, 偶数番目と奇数番目で異なるステップ幅を用いているので, 偶数番目と奇数番目で遷移確率が異なります. そのため, 現在のステップが偶数か奇数かを区別しないと, そもそも遷移確率が定義できません. ここでは, 偶数なら $y = 0$, 奇数なら $y = 1$ となる y を導入して, 状態を (x, y) で指定することにしましょう. マルコフ連鎖であることと既約性が成り立つことはすぐにわかると思います. しかし、y の値は $0, 1, 0, 1, \cdots$ を繰り返すため, 周期は 2 になってしまい, 非周期性が成り立ちません. また、y のことを忘れれば、各ステップで $P(x)T_{c=1}(x \to x') = P(x')T_{c=1}(x' \to x)$, $P(x)T_{c=100}(x \to x') = P(x')T_{c=100}(x' \to x)$ という形の詳細釣り合い条件が成立していますが、y も考慮すると $P(x)T((x,0) \to (x',1)) = P(x')T((x',0) \to (x,1))$, $P(x)T((x,1) \to (x',0)) = P(x')T((x',1) \to (x,0))$ となり、詳細釣り合い条件とは少し異なるものになっています。

そこで, $y = 0 \to 1 \to 0$ という 2 つのステップをまとめて 1 ステップと思うことにしてみま

す. こうするとマルコフ連鎖であることと既約性、非周期性が成り立つことは明らかでしょう. このようにしても詳細釣り合い条件は一般には成り立ちません. [以下略]

4 刷での修正

ページ等	修正前	修正後
p.133, 問題 6.3 の解答	$x(\tau+\Delta\tau) = x(\tau)+\Delta\tau\cdot\frac{dx}{d\tau}(\tau)+\frac{(\Delta\tau)^2}{2}\cdot\frac{dx^2}{d\tau^2}(\tau)+O((\Delta\tau)^3)$	$x(\tau+\Delta\tau) = x(\tau)+\Delta\tau\cdot\frac{dx}{d\tau}(\tau)+\frac{(\Delta\tau)^2}{2}\cdot\frac{d^2x}{d\tau^2}(\tau)+O((\Delta\tau)^3)$
p.191	$D\Phi = F$ を解いて F を求める。	$D\Phi = F$ を用いて F を求める。

5 刷での変更

Chapter 4 に練習問題を追加

<問題 4.7> 連続変数の分布に対して詳細釣り合いを示すとき、ヤコビアンと呼ばれる量が 1 であることを暗黙のうちに用いました。$x \to x'$ という変換で無限小区間 $[x, x+dx]$ が $[x', x', +dx']$ に変化する場合には、幅の変化の割合 $\frac{dx'}{dx}$ がヤコビアンです。この本で扱う例では、特に断りのない限り、ヤコビアンが 1 であることが簡単に示せます。(少々非自明な例に HMC 法があります。問題 6.5 を参照して下さい。) もしヤコビアンが 1 でない場合にはどのような問題が生じ得るでしょうか？

<解答> 確率密度から確率を得るためには、無限小区間 dx を掛ける必要があります。従って、詳細釣り合いの証明に出てきた式には dx や dx' が掛かっていることが暗黙の了解でした。ヤコビアンが 1 であれば、これらは共通の因子であり、無視できました。ヤコビアンが 1 でない場合にはこれらを真面目に取り扱う必要があり、Δx の選び方によっては詳細釣り合いが破れてしまうかもしれません。

Chapter 6 に練習問題を追加

<問題 6.5> リープフロッグ法ではヤコビアンが 1 であることを示してください.

<解答> 表記を簡単にするため, 1 変数の場合を考えます. 多変数の場合もほとんど同じです.

一般に, $(x, p) \to (x', p')$ という変換に伴うヤコビアン J は次のような行列式です:

$$J = \det\begin{pmatrix} \frac{\partial x'}{\partial x} & \frac{\partial p'}{\partial x} \\ \frac{\partial x'}{\partial p} & \frac{\partial p'}{\partial p} \end{pmatrix} = \frac{\partial x'}{\partial x}\frac{\partial p'}{\partial p} - \frac{\partial p'}{\partial x}\frac{\partial x'}{\partial p}. \tag{1}$$

この行列式がリープフロッグ法の各ステップで 1 になっていることは簡単に分かります. $(x,p)\to(x',p')=(x+p\Delta\tau,p)$ というステップでは $J=1\cdot 1-0\cdot\Delta\tau=1$ ですし, $(x,p)\to(x',p')=(x,p-\frac{\partial S}{\partial x}\Delta\tau)$ というステップでは $J=1\cdot 1+\frac{\partial^2 S}{\partial x^2}\Delta\tau\cdot 0=1$ です. リープフロッグ法による時間発展全体のヤコビアンは各ステップのヤコビアンの積なので, これもまた 1 です.

<問題 6.6> 本文中で、HMC 法で $\tau_{\rm fin}=N_\tau\Delta\tau$ を固定したときに N_τ と $\Delta\tau$ の値を調節して効率を上げる方法を説明しました. $\tau_{\rm fin}$ も最適な値に調節するにはどうしたらよいでしょうか？

<解答> N_τ と $\tau_{\rm fin}$ の値を指定すると, 自己相関長 $w(N_\tau,\tau_{\rm fin})$ が評価できます. 独立な配位を一つ得るために必要な計算コストは $N_\tau\times w(N_\tau,\tau_{\rm fin})$ に比例するので、この量が小さくなるように N_τ と $\tau_{\rm fin}$ の値を選びます.

6.1.3 節にコメントを追加

詳細釣り合いの証明の最後、(「他の条件が満たされていることも確認しておきましょう」の直前) に次の 2 文を追加:

上の証明では、ヤコビアンが 1 であることを暗黙のうちに用いています (問題 4.7 参照). リープフロッグ法でヤコビアンが 1 であることは少し計算すれば分かりますので、確認してみてください (問題 6.5).

6刷での修正、変更

ページ等	修正前	修正後
p.11	2012 年に運用が開始された当時世界最先端のスーパーコンピューター「京」は 1 秒間に 1 ペタフロップス $=10^{15}$ 回の浮動小数点計算ができましたが, 単純に足し算だけで良いとしても, $n=10$ だとすでに $100^n/10^{15}=10^5=10$ 万秒, 丸一日以上かかります. $n=12$ だとすでに $100^n/10^{15}=10^9=10$ 億秒, 約 32 年です. 実際には各点での関数の値を計算したりしなければならないのでこの何倍もかかります.	2012 年に運用が開始された当時世界最先端のスーパーコンピューター「京」は 1 秒間に 10 ペタフロップス $=10^{16}$ 回の浮動小数点計算ができましたが, 単純に足し算だけで良いとしても, $n=10$ だとすでに $100^n/10^{16}=10^4=1$ 万秒, 約 3 時間かかります. これを $n=12$ とするだけで, $100^n/10^{16}=10^8=1$ 億秒, 約 3 年です。実際には各点での関数の値を計算したりしなければならないのでこの何倍もかかります.
p.15, Eq. (2.2)	$P_\sigma(x)=\frac{e^{-\frac{(x-\mu)^2}{2\sigma^2}}}{\sqrt{2\pi}\sigma}$	$P_{\sigma,\mu}(x)=\frac{e^{-\frac{(x-\mu)^2}{2\sigma^2}}}{\sqrt{2\pi}\sigma}$
p.131, Eq. (6.61)	$S(x,y)=y^2 f(x)+g(x)$	$S(x,y)=\frac{y^2}{2}f(x)+g(x)$
p.132	すなわち、ガウス乱数 z を生成し	すなわち、分散が 1 のガウス乱数 z を生成し
p.137, 脚注 1	組み合わせの数 $\binom{n}{k}$ は日本の高校数学では ${}_nC_k$ と書かれるのが普通です. 具体的な値は $\binom{n}{k}=\frac{n!}{(n-k)!k!}$ で与えられます.	組み合わせの数 $\binom{n}{k}$ は日本の高校数学では ${}_nC_k$ と書かれるのが普通です. 具体的な値は $\binom{n}{k}=\frac{n!}{(n-k)!k!}$ で与えられます. 表と裏の出る順番も指定した場合にはこの因子は無くなります. いずれにせよ、p には依存しない定数なので、以下では無視します.
p.149	$P(p\|13,7)$	$P(p\|20,13)$
p.151	$\prod_{i=1}^n P(x_i\|\mu,\sigma)$	$\prod_{i=1}^n \rho(x_i\|\mu,\sigma)$
p.156, p.203, p.204, p.206	$e^{-\frac{1}{2}\sum_{i\neq j}\|A_{ij}\|^2-\frac{1}{2}\sum_i\|\mu_i\|^2}$	$e^{-\frac{1}{2}\sum_{i,j}\|A_{ij}\|^2-\frac{1}{2}\sum_i\|\mu_i\|^2}$
p.156	$\Delta S=\frac{1}{2}\sum_{i\neq j}\|A_{ij}\|^2+\frac{1}{2}\sum_i\|\mu_i\|^2$	$\Delta S=\frac{1}{2}\sum_{i,j}\|A_{ij}\|^2+\frac{1}{2}\sum_i\|\mu_i\|^2$
p.157, Fig.7.6	$A_{12}=A21$	$A_{12}=A_{21}$

ページ等	修正前	修正後
p.157	$\frac{1}{2}\sum_{i}^{d}\mu_i^2$	$\frac{1}{2}\sum_{i=1}^{d}\mu_i^2$
p.157	$\frac{1}{2}\sum_{j}^{d}\mu_j^2$	$\frac{1}{2}\sum_{j=1}^{d}\mu_j^2$
p.159	$\Delta E = E(\{s^{(k)}\}) - E(\{s'\})$	$\Delta E = E(\{s'\}) - E(\{s^{(k)}\})$
p.170	n_{cluster} の値を 1 だけ増やし, $i_{n_{\text{cluster}}}$ に追加した格子点の番号を格納する.	$i_{n_{\text{cluster}}}$ に追加した格子点の番号を格納し, n_{cluster} の値を 1 だけ増やす. 最後に, k を 1 だけ増やす.
p.179	$P_1(X) = e^{-f(X)/T_1}$	$P_1(X) \propto e^{-f(X)/T_1}$
p.179	$P_2(X) = e^{-f(X)/T_2}$	$P_2(X) \propto e^{-f(X)/T_2}$
p.180	$\sum_{m=1}^{M}\frac{f[X_m]}{T_m}$	$\sum_{m=1}^{M}\frac{f(X_m)}{T_m}$
p.181, Fig. 7.18	右：$T = 0.001$	右：$T = 0.01$
p.190, Eq. (7.61)	$P(G) = \det(D(G)\cdot D^{\dagger}(G))\cdot e^{-S(G)}$	$P(G) \propto \det(D(G)\cdot D^{\dagger}(G))\cdot e^{-S(G)}$
p.191	計算の大部分は $D\Phi = F$ を解いて F を求めるところと $(DD^{\dagger})\chi = F$ を解いて χ を求めるところに費やされます.	計算の大部分は $(DD^{\dagger})\chi = F$ を解いて χ を求めるところに費やされます.
p.212, Eq. (B.26)	$\frac{e^{-\frac{1}{2}\sum_{i,j=1}^{d}A_{ij}(x_i-\mu_i)(x_j-\mu_j)}}{\sqrt{(2\pi)^d\det A}}$	$\sqrt{\frac{\det A}{(2\pi)^d}}e^{-\frac{1}{2}\sum_{i,j=1}^{d}A_{ij}(x_i-\mu_i)(x_j-\mu_j)}$

図 2.5 の修正

図 2.5 の下側がおかしなものになっている. 正しくは図 1 のようになる.

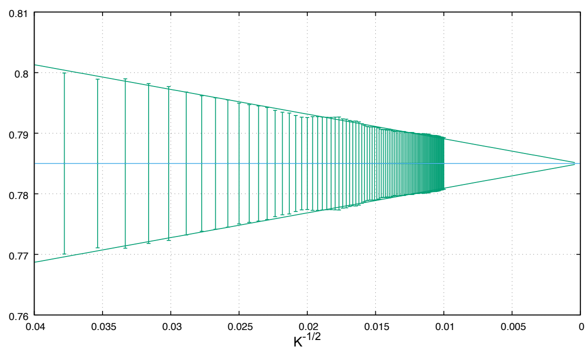

図 1: これが正しい図.