

博士論文

情報復元の統計力学的アプローチ

綴木 潤

2003年2月3日

奈良先端科学技術大学院大学
情報科学研究科 情報処理学専攻

本論文は奈良先端科学技術大学院大学情報科学研究科に
博士(理学) 授与の要件として提出した博士論文である。

綴木 潤

審査委員： 伊藤 実 教授
石井 信 教授
柴田 智広 助教授
岡田 真人 理研 副チームリーダー

情報復元の統計力学的アプローチ*

綴木 潤

内容梗概

雑音のある通信路を経由しての情報伝達の理論（情報理論）は、ほぼ半世紀前に Shannon によって創始されて以来の長い歴史を持っている。情報伝達というと、その通信路にインターネットや無線などをイメージするが、本論文で取り扱う情報伝達経路は広いの意味での通信路であり、スキャナで画像を取り込むまでの過程や音声記憶媒体に記録されるまでの過程など、さまざまな経路による通信路を意味する。これらの通信路において何らかのノイズが付加されることは必然的である。このノイズをいかに削除できるかというのが、情報復元の本質である。

最も簡単な伝達方法として、まず一番最初に思い浮かぶのが、情報を通信路にそのまま送信する方法である。この場合、手がかりになるのは送られてきたノイズ入りの情報のみであるので、我々はこの情報を基に元の情報を推定しなければならない。まず本論文では、この様に符号化なしで情報を送信する場合を扱い、実際に情報を推定することについて議論する。

本論文では符号なしの情報復元問題として、特に画像修復問題を取り扱う。これは通信路に画像を符号化することなく、そのまま伝送する場合に相当する。通信路はスキャナやカメラのレンズ等と見なせる。

劣化画像から元の画像を推定する方法としてベイズ推定を用いる。実は、これまでのベイズ推定に基づく画像修復の枠組では、画像に重畳されるノイズは画素毎に独立であると仮定されていたので、空間的相関を持つようなノイズを取り扱ったモデルは我々の知る限り存在しなかった。しかし光学系の特性等を考慮す

*奈良先端科学技術大学院大学 情報科学研究科 情報処理学専攻 博士論文, NAIST-IS-DT0061019, 2003年2月3日.

れば，重畳されるノイズに空間的相関が発生すると考えるのは自然である．そこで本論文では空間的な相関をもつノイズによって劣化した画像の修復モデルを提案する．送信される画像には何ら冗長性を付加せずに，劣化画像のみからの推定になるので，この場合ノイズを完全に除去することはできない．そこで，提案モデルに対し統計力学的手法を用いることで，どの程度の修復が可能であるかについての定性的な解析を行う．

具体的には原画像の生成モデルとノイズの生成モデルが多次元ガウス分布に従うとし，さらに並進対称な相関行列を仮定することで，両方のモデルにフーリエ変換を用いることを可能にする．このフーリエ変換の利用により，修復誤差等の系の定性的な性質を解析的に計算をすることが可能になる．本論文では，得られた解析結果と人工画像を用いたシミュレーション結果が一致することを示す．さらにハイパーパラメータ推定を行うことで，人工画像と自然画像の画像修復を行う．ハイパーパラメータ推定には修復誤差最小化基準と周辺尤度最大化基準を用いる．そして，既存の空間的相関を持たないノイズモデルがどの程度，空間的相関を持つノイズに対処できるかを見せる．実験として，空間的相関を持つノイズモデルによって生成されたノイズを，あえて既存の空間的相関を持たないノイズモデルを用いて修復を行った．そして既存の空間的相関を持たないノイズモデルでは，空間的相関を持つノイズをうまく修復できないことを示す．特に周辺尤度最大化を用いる場合は，上記のように生成用のノイズモデルと修復用のノイズモデルが一致していなければ，修復が非常に困難になる様子を見せる．

さらに，情報をそのまま送信する方法とは異なり，ノイズを後で除去する手がかりを付加することで，情報を冗長にしてから送信する符号化について述べる．冗長性をうまく利用することで元の情報に復元することができる．この符号化の代表的なものに Sourslas 符号がある．Sourslas 符号とはデータ 1 ビットをスピンと見なし，それらの積で相互作用を作るものであり，Hopfield による連想記憶モデルと深い関係がある．連想記憶モデルでは，ハミルトニアン¹の極小が記憶パターンに対応し，記憶パターンへの緩和過程としてとらえられるが，Sourslas はこの Hopfield の仕事に影響を受け，誤り訂正符号のスピンモデルを提案した．Sourslas 符号でも，符号の誤りを除いた状態がハミルトニアン¹の極小に対応している．実

は画像修復も，Hopfield 型の連想記憶モデルや Sourlas 符号と同様に，ハミルトニアン基底状態探索問題に置き換えることができる．情報復元の失敗という点において Sourlas 符号と連想記憶モデルは類似している．Sourlas 符号の失敗とは復号化の失敗であるが，連想記憶モデルの場合は偽記憶と呼ばれる記憶パターンではない記憶に収束することである．

本論文ではさらに連想記憶モデルの欠点である，偽記憶に収束しにくいモデルとしてハイブリッド型連想記憶モデルを提案する．提案モデルでは，パターンの表記部（可視部）以外に，連想処理専用の隠れ部を付け加えている．可視部と隠れ部の写像に線形写像を用いることにより，可視部 隠れ部 可視部の一連の再帰的結合を，2層の階層型パーセプトロンと見なすことができる．可視部と隠れ部の結合の学習には誤差逆伝播法を用いる．一方で，可視部の相互結合にはヘップ則を用いる．本論文で用いる誤差逆伝播法は，可視部と隠れ部の結合の学習において記憶パターンのみを用いるので，その誤差逆伝播法を教師なし学習と見なせる．可視部の学習則であるヘップ則も教師なし学習であることから，使用した学習法はハイブリッド型の教師なし学習である．

一般に，連想記憶モデルでは多数の偽記憶に収束する現象が生じうるが，本論文では，提案するハイブリッド型モデルがより広い引き込み領域を持ち，偽記憶に収束しにくいモデルであることをシミュレーション結果により示す．

キーワード

情報復元，情報伝達，符号化，復号化，Sourlas 符号，画像修復，統計力学，ベイズ推定，ハイパーパラメータ推定，フーリエ変換，神経回路，連想記憶，ハイブリッド型学習

研究業績リスト

論文

- Jun Tsuzurugi, Masato Okada:
Statistical Mechanics of Bayesian Image Restoration under Spatially Correlated Noise,
Physical Review E, Vol 66, 066704, Dec, 2002, (3章の内容)
- 綴木 馴, 高橋 規一, 石井 信:
ハイブリッド型学習による隠れ素子付き連想記憶モデル,
システム制御情報学会論文誌 15(11),2002, (4章の内容)

査読付き国際会議

- Jun Tsuzurugi, Masato Okada:
Bayesian Image Restoration under Spatially Correlated Noise -Statistical mechanical approach-, *Proceeding for 9th International Conference on Neural Information Processing*, (ICONIP2002)

研究会

- 綴木 馴, 岡田真人:
空間的な相関を持つノイズ下での画像修復, ニューロコンピューティング研究会, 電子情報通信学会技術研究報告, NC2001-67, (2001-11)
- 綴木 潤, 高橋規一:
隠れ素子付き相互結合型ネットワークにヘップ則とBP法を適用した自己連想記憶モデル, ニューロコンピューティング研究会, 電子情報通信学会技術研究報告, NC98-69, (2000-03)
- 綴木 潤, 高橋規一:
複数のホップフィールド型ネットワークを用いた連想記憶方式の提案, ニュー

ロコンピューティング研究会, 電子情報通信学会技術研究報告, NC98-69,
(1999-01)

物理学会

- 綴木 馴, 岡田真人:
空間的な相関を持つノイズ下での画像修復, 日本物理学会分科会, 2001 年秋
季大会, (18pRD-15)

国内発表

- 綴木 潤, 高橋 規一:
可視部において力学系を形成する隠れ素子付き連想記憶モデル, 平成 11 年
度電気関係学会, 九州支部連合会大会論文集, (1999)

備考:

上記で掲載されている名前, 綴木 馴(つづるぎ じゅん)は著者が 2001 年 4 月か
らペンネームとして用いた名前で, 著者の本名である綴木 潤(つづるぎ じゅん)
と同一人物である.

目次

第1章	論文の構成	1
第2章	これまでの先行研究	5
2.1.	ガウスモデルによる画像修復	5
2.1.1	ガウスモデルによる画像の生成確率	5
2.1.2	ガウスモデルによる画像の劣化過程	7
2.1.3	画像修復の定式化	8
2.1.4	本節のまとめ	9
2.2.	Sourlas 符号による誤り訂正	11
2.2.1	もっとも簡単な符号化とパリティ検査符号	11
2.2.2	2体相互作用の Sourlas 符号	12
2.2.3	一般の Sourlas 符号	14
2.2.4	最大事後確率法とハミルトニアン	15
2.2.5	有限温度復号	17
2.2.6	本節のまとめ	17
2.3.	連想記憶モデル	19
2.3.1	相互結合型ネットワーク	20
2.3.2	ヘップ則による学習	21
2.3.3	記憶パターンの想起	23
2.3.4	連想記憶モデルのハミルトニアン	25
2.3.5	連想記憶モデルの性質とシミュレーション	26
2.3.6	本節のまとめ	34

第 3 章	画像修復の統計力学	35
3.1.	この章のはじめに	35
3.2.	モデル	37
3.2.1	画像の生成過程	37
3.2.2	画像の劣化過程	39
3.3.	ベイズ推定と画像修復の有効ハミルトニアン	40
3.4.	理論	41
3.4.1	修復画像の期待値	41
3.4.2	修復誤差最小化基準	44
3.4.3	周辺尤度最大化基準	46
3.5.	シミュレーション	48
3.5.1	修復誤差最小化基準	48
3.5.2	周辺尤度最大化基準	51
3.5.3	サンプル画像	54
3.5.4	ノイズモデルの特異点について	56
3.6.	議論	59
3.7.	本章のむすび	61
3.8.	本章の付録	63
3.8.1	k について	63
3.8.2	式 (3.20) の導出法	64
第 4 章	ハイブリッド型連想記憶モデル	67
4.1.	この章のはじめに	67
4.2.	ネットワークの定式化	69
4.3.	隠れ素子付き相互結合型ネットワーク	69
4.3.1	モデルの拡張	71
4.3.2	可視空間から隠れ空間への写像	72
4.4.	連想記憶モデルの構築	73
4.4.1	学習法	75
4.4.2	離散型	77

4.5.	シミュレーション	78
4.5.1	偽記憶への収束率	78
4.5.2	臨界オーバーラップ	80
4.6.	本章のむすび	82
4.7.	付録：誤差逆伝播法	83
第5章	総括	87
	謝辞	90
	参考文献	92

目次

1.1	情報復元の図．元の情報 ξ を通信路に送信することで，劣化した情報 τ が得られる． τ から元の情報を推定することで復元情報 σ を求める．	1
1.2	冗長化による符号化と復号化の図．元の情報 ξ を冗長化することで符号化した情報を J^{in} とし，通信路に送信する．通信路から得られた情報 J^{out} を復号化することで，復元情報 σ を求める．	2
1.3	本論文の内容の流れをあらわす図．	3
2.1	相互結合型ネットワークの概略図	21
2.2	符号関数	22
2.3	素子数 $N = 1000$ ，パターン数 $P = 120$ ， $(\alpha = 0.12)$ のときの想起の様子．横軸が時刻 t をあらわし，縦軸が各時刻でのオーバーラップ $m(t)$ をあらわす．	26
2.4	素子数 $N = 1000$ ，パターン数 $P = 200$ ， $(\alpha = 0.20)$ のときの想起の様子．縦軸が時刻 t をあらわし，横軸が各時刻でのオーバーラップ $m(t)$ をあらわす．	27
2.5	連想記憶モデルの相転移をあらわすグラフ．横軸は記憶率 α ，縦軸は系が平衡状態に達したときのオーバーラップ $m(\infty)$ をあらわす．初期状態を $m(0) = 1$ とした場合において平衡状態に達したときのオーバーラップ $m(\infty)$ を見た．実線は 11 回のシミュレーションにより得られたデータの平均をあらわし，エラーバーは標準偏差をあらわす．	28

2.6	素子数 $N = 320$, 様々な初期オーバーラップ $m(0)$ から開始した場合における想起の成功率 . 横軸が記憶パターン数 P をあらわし , 縦軸が想起の成功率をあらわす	31
2.7	記憶パターン自身を初期状態にして ($m(0) = 1$) 想起の成功率をあらわすグラフ . 縦軸が想起の成功率 , 横軸が記憶率 α をあらわす	32
2.8	双曲正接	33
3.1	原画像との平均二乗誤差 . 横軸が式 (3.11) の a であり , 縦軸が原画像との平均二乗誤差をあらわす . (A). 式 (3.52) で与えられる最適なデコードをした場合 . (B). 修復誤差最小化により空間的な相関が無いモデルでデコードした場合 . (C). 周辺尤度最大化により \hat{r} を推定した場合 . (D). 周辺尤度最大化により $\hat{\beta}, \hat{h}, \hat{r}$ を推定した場合 . (E). 式 (3.54) で与えられるデコード前の誤差	49
3.2	\hat{r} の推定結果をあらわすグラフ . 横軸は式 (3.11) の a であり , 縦軸は推定値 \hat{r} の平均値と修復誤差最小化により得られた最適値 \hat{r} をあらわす . 実線は図 3.1 の (B) に用いた \hat{r} , 波線は図 3.1 の (C) に用いた \hat{r} , 実線は図 3.1 の (D) に用いた \hat{r} である	50
3.3	図 3.1 の (D) で用いた \hat{h} (点線) , $\hat{\beta}$ (波線) の平均値をあらわすグラフ . 横軸は式 (3.11) の a であり , 縦軸は推定値 $\hat{\beta}, \hat{h}$ の平均値と正しい値 β, h をあらわす . 実線は正しい値 $h = \beta = 1.0$ である	52
3.4	平均二乗誤差と対数周辺尤度のグラフ . 横軸はハイパーパラメータ \hat{r} をあらわし , 左の縦軸は式 (3.50) で与えられる平均二乗誤差 (B) あらわし , 右の縦軸は式 (3.63) で与えられる対数周辺尤度 (C) をあらわす . 用いたパラメータは $a = 1.0, \beta = 1.0, h = 1.0, \kappa = 3.0$ とした . (B) で用いた条件は図 3.1(B) の $a = 1.0$ における条件と同一であり , 同様に , (C) で用いた条件は , は図 3.1(C) の $a = 1.0$ における条件と同一である	53
3.5	人工画像 . $N = 64^2$	54
3.6	自然画像 . $N = 64^2$	55

3.7	元の値を $a = 0.5, b = 0.75, \kappa^2 = 9.0$ とした場合に, 漸化式 (3.69), (3.70) において $\hat{a}, \hat{\kappa}^2$ を繰り返し演算した場合の相図. 横軸に $\hat{\kappa}^2$, 縦軸に \hat{a} としている. $\hat{b} = 0.75$ と固定した. 正しい値 $\hat{a} = 0.5, \hat{\kappa}^2 = 9.0$ を代入した場合でも, \hat{a} がその定義域を超えて発散してしまう様子がわかる.	58
3.8	原画像と修復画像の平均二乗誤差. 横軸は式 (3.11) の a を示し, 縦軸は平均二乗誤差をあらわす. 式 (3.52) で与えられる修復誤差最小化基準を用いてデコードしている. パラメータ $b = 0.75$ と $N = 64^2$ は固定している. 実線 (i) では, $\beta = 1.0, h = 1.0, \kappa = 3.0$ を用い, 点線 (ii) では, $\beta = 0.5, h = 10^{-4}, \kappa = 7.0$ を用いている.	59
3.9	図 3.8(ii) の典型例. $N = 64^2$ とした. 原画像で用いたパラメータは $\beta = 0.5$ と $h = 10^{-4}$ であり, 3つの原画像は同一である. ノイズ画像で用いられたパラメータは $\beta = 0.5, h = 10^{-4}, \kappa = 7.0, b = 0.75$ であり, a のみが増加している. $a = 0.55$ のとき, 原画像とノイズ画像が見た目に似ており, 修復が最もうまく行っていないのが分かる.	60
3.10	様々な κ に対するノイズ画像. $a = 0.0, b = 0.75$ と固定し, κ を $\kappa = 0.5, 3.0, 7.0$ と変化させた場合.	63
4.1	隠れ部の導入の様子をあらわす図.	70
4.2	状態変数の軌道. 可視素子の数 $N = 2$, 隠れ素子の数 $R = 1$. 可視素子 1 と可視素子 2 の可視変数が動く部分空間が可視空間 (2次元) になり, 隠れ素子 1 の隠れ変数が動く部分空間が隠れ空間 (1次元) になる.	71
4.3	式 (4.26) の第三項 [] 内 $B\{g^Y(Ax + \theta) + \phi\}$ の展開図. 入力層の i 番目の素子と中間層の j 番目の素子の結合荷重は A_{ij} , 中間層の j 番目の素子と出力層の k 番目の素子の結合荷重は B_{kj} , 中間層の j 番目の素子にかかるバイアスは θ_j , 出力層の k 番目の素子にかかるバイアスは $B_{kR+1}\phi_k$ となる. これらの結合線は分かりやすいように他の線と区別して太線を用いている.	76

4.4	横軸に a , 縦軸に偽記憶への収束回数をプロットした図	77
4.5	連続型における出力パターンと記憶パターンのオーバーラップをあらわすグラフ	79
4.6	離散型における出力パターンと記憶パターンのオーバーラップをあらわすグラフ	80
4.7	臨界オーバーラップ . 連続型	81
4.8	臨界オーバーラップ . 離散型	82
4.9	階層型ネットワークの概念図	83

第1章 論文の構成

ある場所から別のある場所に N ビット列のベクトルで表現されている情報 ξ を送ることを考える．この情報伝送のための経路を通信路 (channel) と呼ぶ．通信路には何らかのノイズがあり，送信された情報と通信路からの出力とは一般に異なるのが普通である．本論文ではこのようにノイズの混じった出力情報から元の情報を求める情報復元について議論する．

通信路のノイズによって劣化した情報を，元の情報に復元する方法として，まず考えられるものは，劣化した情報を基に元の情報を推定する方法である．この場合，我々は元の情報 ξ をそのまま通信路に送信し，出力として劣化した情報 τ を得る．この劣化情報には何らかの形でノイズが入っているので， τ から元の情報を推定する必要がある．推定された情報 σ を復元情報とする．この様子を図 1.1 に示す．

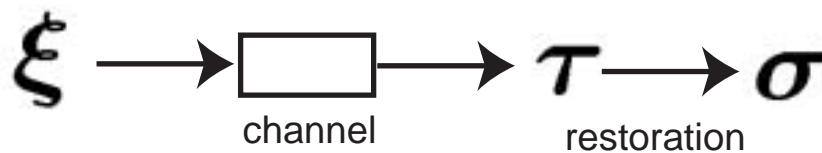


図 1.1 情報復元の図．元の情報 ξ を通信路に送信することで，劣化した情報 τ が得られる． τ から元の情報を推定することで復元情報 σ を求める．

次に，より高度な方法として，元の情報を求めるための手がかりをあらかじめ付加するために，情報を冗長にしてから送信する符号化について考える．符号化された情報はノイズが混ざった状態で通信路から出力されるが，この冗長性をうまく利用し復号化を行うことで，復元情報を求めることができる．この様子を図 1.2 に示す．劣化した情報から元の情報を推定するだけでは，出力ビットのどの部分が間違っていて，どの部分が正しいのか分からないが，情報を符号化した場

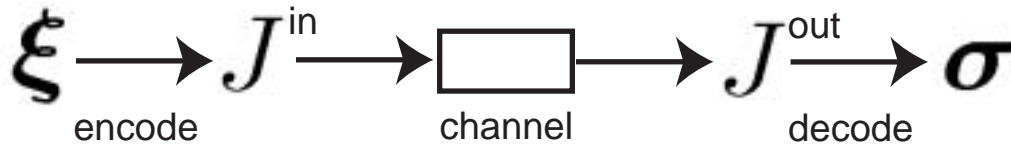


図 1.2 冗長化による符号化と復号化の図．元の情報 ξ を冗長化することで符号化した情報を J^{in} とし，通信路に送信する．通信路から得られた情報 J^{out} を復号化することで，復元情報 σ を求める．

合は，復号化を行うことで完全に元と同じ情報を求めることができる場合が多いので，どの部分が間違っていたのかも分かる．

このように本論文では，符号化がなされない場合の情報復元と，符号化がなされた場合の情報復元の両方について述べる．本論文の構成を簡単にイラスト化したものを図 1.3 に示す．

一般に符号化しない場合は完全な情報が得られないので，通信路に送信するときは，なるべく正確に情報を復元するためにも符号化することが望まれる．しかし，実際には符号化を施せない場合が存在しうる．例えば，スキャナやカメラで画像を撮影する場合やマイクで音声を録音する場合である．このような場合，レンズやマイクに送信する前にあらかじめ情報を符号化することができない．したがって，ノイズを除去しようと思えば，伝達されてきたノイズ入りの情報のみで元の情報を復元しなくてはならない．

図 1.3 に示すように本論文ではまず，符号化のない場合の情報復元 (Information restoration) について述べる．特に，情報復元の中でも画像修復 (Image restoration) に焦点を絞って議論する．まずは 2.1 節において先行研究であるガウスモデルによる画像修復について解説する．さらに第 3 章では，統計力学的手法を用いることで，どの程度の画像修復が可能であるか等画像修復の定性的な性質について議論する．

画像修復は冒頭で述べたとおり，符号化することなく情報を通信路に送信することに相当する．したがって，我々はノイズの入った劣化画像から元の画像を推定する必要がある．推定法にはベイズ推定を用いる．既存のベイズ推定を用いた画像修復問題では，空間的相関を持ったノイズモデルは提案されていなかったが，

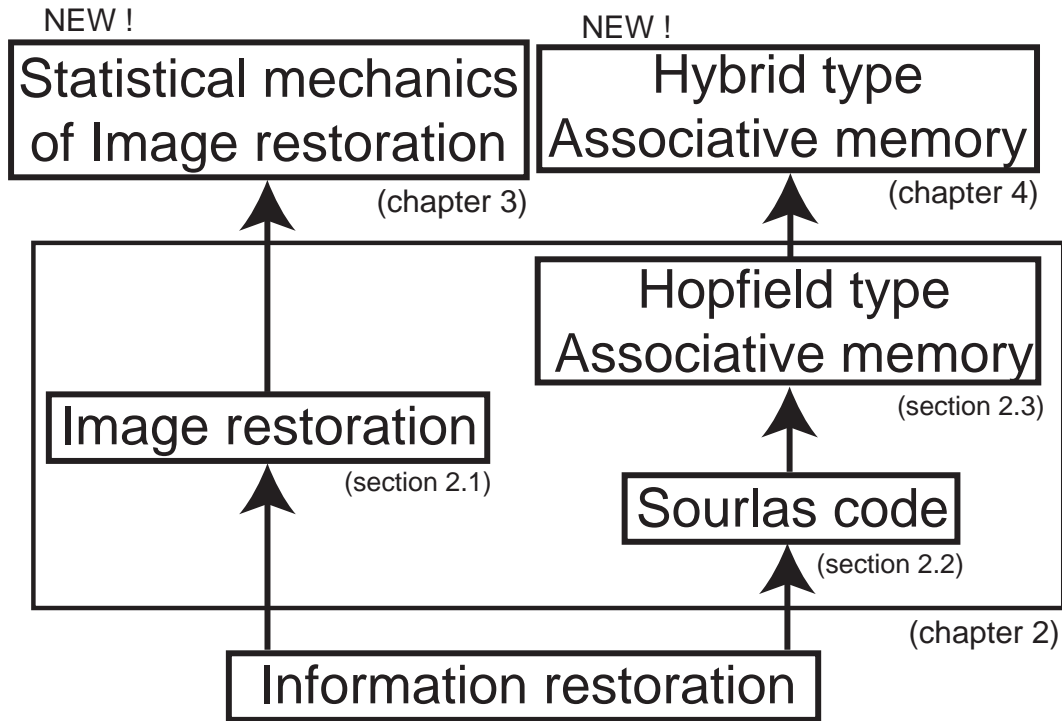


図 1.3 本論文の内容の流れをあらわす図 .

第 3 章では新たに空間的な相関を持ったノイズモデルを提案する．さらに提案モデルの統計力学的な解析を行う．そして，人工画像や自然画像において数値計算を行うことで，解析結果との比較を行う．

次に 2.2 節では，符号化された場合の情報復元について述べる．符号の中でも特に Sourlas 符号 (Sourlas code) を取り扱い，統計力学的な手法について紹介する．図 1.3 に示すように，Hopfield 型の連想記憶モデル (Hopfield type Associative memory)[37] は Sourlas 符号 [39] の拡張モデルとなっている．歴史的には Hopfield 型の連想記憶モデルの方が古く，実際に Sourlas は Hopfield の影響を多分に受けているのだが，エネルギー緩和過程という点では，Hopfield 型の連想記憶モデルの方が Sourlas 符号の拡張モデルになっていると言える．2.3 節では連想記憶モデルについての統計力学的な扱い紹介するとともに，シミュレーションを用いることで連想記憶モデルの性質を見ていく．実は，画像修復問題も，Sourlas 符号も，Hopfield 型の連想記憶モデルも，ハミルトニアンの基底状態探索問題に置き換え

ることができる．このことについては第 2 章において説明する．

ところで，2.3 節で述べる連想記憶モデルには偽記憶に収束するという欠点がある．この偽記憶問題を緩和するため，第 4 章ではさらに，ヘブ則と誤差逆伝播法の二つの学習則を用いたハイブリッド型の連想記憶モデルを提案する．モデルの能力の向上を期待して，パターン表記部である可視部とは別に連想処理専用の隠れ素子を導入する．さらに可視部と隠れ部の関係に線形写像を用いることで，ヘブ則と誤差逆伝播法の二つの学習則を実装する方法を提案する．これら二つの学習則が効果的に作用し，結果的に記憶パターンの引き込み領域が広くなることをシミュレーションによって示す．

第2章 これまでの先行研究

本章では，先行研究である，ガウスモデルによる画像修復，Sourlas 符号，連想記憶モデルについて述べる．画像修復も，Sourlas 符号や連想記憶モデルと同様に，ハミルトニアン基底状態探索問題に置き換えることができる．このことについても説明する．

2.1. ガウスモデルによる画像修復

2.2 節で述べるように，符号化と復号化を用いることで元のデータを完全に復元することが可能であるので，情報を送信するときは符号化することが望ましいが，実際には符号化できない場合が存在しうる．例えば，スキャナやカメラで画像を撮影する場合やマイクで音声を録音する場合は，情報をレンズやマイクに送信する前にあらかじめ符号化することができない．よって，情報は劣化過程である通信路に送信されることになる．このような場合においてノイズを除去しようとする場合は，伝達されてきた情報のみで元の情報を推定しなくてはならない．そこで本節では画像修復問題に焦点を当て，まずは符号化しない場合の情報修復について議論する．本節では先行研究であるガウスモデルによる画像修復 [16] について解説する．

2.1.1 ガウスモデルによる画像の生成確率

画像は一辺あたりの画素数を L と正方形で与えられるとする．この場合，素子の総数 N は $N = L^2$ となる．格子点 i 上の原画像の画素値を ξ_i とおき，原画像をベクトル $\xi = \{\xi_i\}$ であらわす．原画像は 2 次元ガウス分布にしたがって生成さ

れると仮定する .

$$P(\boldsymbol{\xi}) = \frac{1}{Z_{\text{prior}}(\beta, h)} \exp[-H(\boldsymbol{\xi}, \beta, h)] \quad (2.1)$$

ただし , β, h は正のスカラーである . $Z_{\text{prior}}(\beta, h)$ は規格化定数である . 式 (2.1) の H は統計力学でいうハミルトニアンであり , 次で与えられるとする .

$$H = \beta \sum_{\langle ij \rangle} (\xi_i - \xi_j)^2 + h \sum_j \xi_j \quad (2.2)$$

ξ_i は連続実数変数である . $\beta, h > 0$ とする . 右辺第 1 項の和は 2 次元最近接格子点であるが , 並進対称な範囲であれば拡張は容易である .

フーリエ変換で解くので , 変換の定義を次のように行う .

$$\tilde{\xi}_{\mathbf{k}} = \frac{1}{\sqrt{N}} \sum_j \xi_j e^{-i\mathbf{k} \cdot \mathbf{j}} \quad (2.3)$$

逆フーリエ変換 (フーリエ表記) を次で定義する .

$$\xi_j = \frac{1}{\sqrt{N}} \sum_{\mathbf{k}} \tilde{\xi}_{\mathbf{k}} e^{i\mathbf{k} \cdot \mathbf{j}} \quad (2.4)$$

ここで $\mathbf{k} \cdot \mathbf{j}$ はこれらのベクトルの積をあらわしている . i は虚数単位である . \mathbf{k} も \mathbf{j} と同じ 2 次元ベクトルであり , 自由度は $N = L^2$ である . また , \mathbf{k} の各成分 k_x, k_y は以下のような値をとる .

$$0, \frac{2}{L}\pi, \frac{4}{L}\pi, \dots, \frac{2(L-1)}{L}\pi \quad (2.5)$$

ハミルトニアン (2.2) のフーリエ表記は

$$H = \sum_{\mathbf{k}} (\beta \tilde{G}_{\mathbf{k}} + h) \tilde{\xi}_{\mathbf{k}} \tilde{\xi}_{-\mathbf{k}} \quad (2.6)$$

となる . ここで $\tilde{G}_{\mathbf{k}}$ は

$$\tilde{G}_{\mathbf{k}} = 2(2 - \cos k_x - \cos k_y) \quad (2.7)$$

である . ここに来て $Z_{\text{prior}}(\beta, h)$ も次のように与えられる .

$$Z_{\text{prior}}(\beta, h) = (2\pi)^{\frac{N}{2}} \left| 2(\beta \tilde{G}_{\mathbf{k}} + h) \right|^{-\frac{1}{2}} \quad (2.8)$$

$\tilde{\xi}_{\mathbf{k}}$ を実部と虚部に分けて $\tilde{\xi}_{\mathbf{k}} = \tilde{\xi}_{\mathbf{k}}^{(1)} + i\tilde{\xi}_{\mathbf{k}}^{(2)}$ と書くことにして, \mathbf{k} の和を \mathbf{k} と $-\mathbf{k}$ についての和の $1/2$ で置き換えれば

$$\begin{aligned} H &= \sum_{\mathbf{k}} (\beta\tilde{G}_{\mathbf{k}} + h) \frac{1}{2} (\tilde{\xi}_{\mathbf{k}}\tilde{\xi}_{-\mathbf{k}} + \tilde{\xi}_{-\mathbf{k}}\tilde{\xi}_{\mathbf{k}}) \\ &= \sum_{\mathbf{k}} (\beta\tilde{G}_{\mathbf{k}} + h) \{(\tilde{\xi}_{\mathbf{k}}^{(1)})^2 + (\tilde{\xi}_{\mathbf{k}}^{(2)})^2\} \end{aligned} \quad (2.9)$$

となる. ここで $\tilde{\xi}_{\mathbf{k}}$ と $\tilde{\xi}_{-\mathbf{k}}$ が複素共役であることを使った. 式 (2.9) は, 実部と虚部が独立な変数として振る舞うことを示している.

2.1.2 ガウス模型による画像の劣化過程

2次元ベクトル $\tau = \{\tau_i\}$ を劣化画像する. 画像の劣化過程は次のようにあらわされる.

$$P_{\text{out}}(\tau|\xi) = \frac{1}{(\sqrt{2\pi v})^N} \exp \left\{ -\frac{1}{2v^2} \sum_{\mathbf{i}} (\tau_{\mathbf{i}} - \xi_{\mathbf{i}})^2 \right\} \quad (2.10)$$

フーリエ表記すると,

$$\begin{aligned} P_{\text{out}}(\tau|\xi) &= \frac{1}{(\sqrt{2\pi v})^N} \exp \left\{ -\frac{1}{2v^2} \sum_{\mathbf{k}} (\tilde{\tau}_{\mathbf{k}} - \tilde{\xi}_{\mathbf{k}})(\tilde{\tau}_{-\mathbf{k}} - \tilde{\xi}_{-\mathbf{k}}) \right\} \\ &= \frac{1}{(\sqrt{2\pi v})^N} \exp \left\{ -\frac{1}{2v^2} \sum_{\mathbf{k}} (\tilde{\tau}_{\mathbf{k}}^{(1)} - \tilde{\xi}_{\mathbf{k}}^{(1)})^2 - \frac{1}{2v^2} \sum_{\mathbf{k}} (\tilde{\tau}_{\mathbf{k}}^{(2)} - \tilde{\xi}_{\mathbf{k}}^{(2)})^2 \right\} \end{aligned} \quad (2.11)$$

となる. 最後の表式は前節と同様にして実部と虚部に分けることにより得られる. v^2 はガウス分布の分散である. この場合, 重畳されるノイズは空間的な相関を持たない.

2.1.3 画像修復の定式化

ガウス模型を使った画像修復を定式化しよう．ベイズの公式により次が成り立つ．

$$\begin{aligned} P(\sigma|\tau) &= \frac{P_{\text{out}}(\tau|\sigma)P(\sigma)}{\int d\sigma P_{\text{out}}(\tau|\sigma)P(\sigma)} \\ &= \frac{\exp[-H_{\text{eff}}]}{\int d\sigma \exp[-H_{\text{eff}}]} \end{aligned} \quad (2.12)$$

ただし，

$$H_{\text{eff}} = \sum_{\mathbf{k}} (\hat{\beta}\tilde{G}_{\mathbf{k}} + \hat{h})(\tilde{\sigma}_{\mathbf{k}}^{(1)})^2 + \frac{1}{2\hat{v}^2} \sum_{\mathbf{k}} (\tilde{\tau}_{\mathbf{k}}^{(1)} - \tilde{\sigma}_{\mathbf{k}}^{(1)})^2 + (1 \rightarrow 2) \quad (2.13)$$

であり， $\hat{\beta}, \hat{h}, \hat{v}$ はそれぞれ β, h, v の推定値である．ベイズ推定では $P(\sigma|\tau)$ を事後分布， $P_{\text{out}}(\tau|\sigma)$ を尤度， $P(\sigma)$ を事前分布と呼ぶ．さらに式 (2.13) で与えられる H_{eff} を統計力学では有効ハミルトニアンと呼ばれる．

この確率分布による期待値 $\langle \sigma_i \rangle$ を修復結果とする．よりもっともらしい $\langle \sigma_i \rangle$ を求めるには事後確率 $P(\sigma|\tau)$ を最大化するればよく，このためには式 (2.12) の右辺の分子を最大化すれば良い．すなわち，有効ハミルトニアン (2.13) を最小化すれば良いことになる．つまり，よりもっともらしい修復画像 σ を求めることは有効ハミルトニアン (2.13) の基底状態探索問題に帰着される．このような方法を最大事後確率 (Maximum A Posteriori, MAP) 法と呼ぶ．

修復結果を具体的に計算する．

$$\langle \sigma_j \rangle = \sum_{\mathbf{k}} \langle \tilde{\sigma}_{\mathbf{k}}^{(1)} + i\tilde{\sigma}_{\mathbf{k}}^{(2)} \rangle e^{i\mathbf{k}\cdot\mathbf{j}} \quad (2.14)$$

であるから，フーリエ成分の期待値が分かればよい．

$$\langle \tilde{\sigma}_{\mathbf{k}}^{(1)} \rangle = \frac{\int \prod_{\mathbf{k}} d\tilde{\sigma}_{\mathbf{k}} \tilde{\sigma}_{\mathbf{k}}^{(1)} e^{-H_{\text{eff}}}}{\int \prod_{\mathbf{k}} d\tilde{\sigma}_{\mathbf{k}} e^{-H_{\text{eff}}}} \quad (2.15)$$

であり， H_{eff} も各フーリエ成分に分かれるから，ひとつの k のみに着目して積分を実行すれば十分である．また，実部 (上付き (1) の変数) のみ考えればよい．

虚部は分子分母でうち消す．上式分母で k の部分のみの寄与を $Z_{\mathbf{k}}$ とすれば

$$\begin{aligned}\langle \tilde{\sigma}_{\mathbf{k}}^{(1)} \rangle &= \frac{1}{Z_{\mathbf{k}}} \int d\tilde{\sigma}_{\mathbf{k}}^{(1)} \tilde{\sigma}_{\mathbf{k}}^{(1)} \exp \left\{ -(\hat{\beta}\tilde{G}_{\mathbf{k}} + \hat{h} + (2\hat{v}^2)^{-1})(\tilde{\sigma}_{\mathbf{k}}^{(1)})^2 + \frac{1}{\hat{v}^2}\tilde{\sigma}_{\mathbf{k}}^{(1)}\tilde{\tau}_{\mathbf{k}}^{(1)} + (\tilde{\tau}_{\mathbf{k}}^{(1)} \text{の項}) \right\} \\ &= \frac{1}{Z_{\mathbf{k}}} \int d\tilde{\sigma}_{\mathbf{k}}^{(1)} \tilde{\sigma}_{\mathbf{k}}^{(1)} \exp \left\{ -A_{\mathbf{k}}(\tilde{\sigma}_{\mathbf{k}}^{(1)} - \frac{B\tilde{\tau}_{\mathbf{k}}^{(1)}}{A_{\mathbf{k}}}) + (\tilde{\tau}_{\mathbf{k}}^{(1)} \text{の項}) \right\} = \frac{B\tilde{\tau}_{\mathbf{k}}^{(1)}}{A_{\mathbf{k}}}\end{aligned}\quad (2.16)$$

が得られる．ここで

$$A_{\mathbf{k}} = \hat{\beta}\tilde{G}_{\mathbf{k}} + \hat{h} + (2\hat{v}^2)^{-1}, \quad B = \frac{1}{2\hat{v}^2}\quad (2.17)$$

でありまた，最後の行を導く際には分子分母で $\tilde{\tau}_{\mathbf{k}}^{(1)}$ の項が打ち消し合うことを使った．全く同様の式が虚部についても成立する．それゆえ

$$\begin{aligned}\langle \sigma_{\mathbf{j}} \rangle &= \frac{B}{\sqrt{N}} \sum_{\mathbf{k}} \frac{1}{A_{\mathbf{k}}} (\tilde{\tau}_{\mathbf{k}}^{(1)} + i\tilde{\tau}_{\mathbf{k}}^{(2)}) e^{i\mathbf{k}\cdot\mathbf{j}} \\ &= \frac{1}{\sqrt{N}} \sum_{\mathbf{k}} \frac{B\tilde{\tau}_{\mathbf{k}}}{A_{\mathbf{k}}} e^{i\mathbf{k}\cdot\mathbf{j}} \\ &= \frac{1}{\sqrt{N}\hat{v}^2} \sum_{\mathbf{k}} \frac{\sum_{\mathbf{i}} \tau_{\mathbf{i}} e^{-i\mathbf{k}\cdot\mathbf{i} + i\mathbf{k}\cdot\mathbf{j}}}{\hat{\beta}\tilde{G}_{\mathbf{k}} + \hat{h} + (2\hat{v}^2)^{-1}} \\ &= \frac{1}{N\hat{v}^2} \sum_{\mathbf{i}} \tau_{\mathbf{i}} \sum_{\mathbf{k}} \frac{e^{i\mathbf{k}\cdot(\mathbf{j}-\mathbf{i})}}{\hat{\beta}\tilde{G}_{\mathbf{k}} + \hat{h} + (2\hat{v}^2)^{-1}} \\ &= \frac{1}{N\hat{v}^2} \sum_{\mathbf{i}} \tau_{\mathbf{i}} \sum_{\mathbf{k}} \frac{\cos \mathbf{k}\cdot(\mathbf{j}-\mathbf{i})}{\hat{\beta}\tilde{G}_{\mathbf{k}} + \hat{h} + (2\hat{v}^2)^{-1}}\end{aligned}\quad (2.18)$$

上式は劣化画像 τ の線形変換で修復画像 $\sigma = \{\sigma_{\mathbf{i}}\}$ が直ちに求まることを示している．このように，フーリエ変換を用いることで修復結果を簡単に求めることができる．

2.1.4 本節のまとめ

本節では符号化しない場合の情報修復の一例として，ガウス模型を用いた場合の画像修復について紹介した．フーリエ変換を用いることで修復結果が簡単に求まることを示した．また，画像修復問題が有効ハミルトニアンの基底状態探索問題に帰着できることを紹介した．

第3章ではさらに空間的な相関を持つノイズを修復するモデルについて議論する。本節では、有効ハミルトニアンを最小化すればもっともらしい修復結果が得られることは示したものの、どのような場合に修復誤差が最小になるかについては議論していない。第3章ではこのような議論を統計力学的な立場から行うことでモデルの定性的な評価を行う。さらに本節では紹介しなかったハイパーパラメータ推定も行うことで実際の画像修復も行い、本節で紹介した空間的な相関を持たないモデルが、空間的な相関をもつノイズにどれくらい対応できるかについても議論する。

2.2. Sourlas 符号による誤り訂正

本節では、符号化の概念とその統計力学的な取り扱いを Sourlas 符号 [39] を通して紹介する。

ある場所から別のある場所に何らかの情報を送ることを考える。この情報伝送のための経路を通信路 (channel) と呼ぶ。長さ N ビット列で表現されているベクトル ξ を伝送する情報とする。通信路には何らかのノイズがあり、通信路からの出力は送信された情報とは一般に異なるのが普通である。ノイズの混じった出力から元の情報を求めることが情報復号である。

元の情報をそのまま通信路に送信したのでは、出力ビットのどの部分が間違っていて、どの部分が正しいのか分からない。そこで、元の情報を求めるための手がかりを付加し情報を冗長 (redundant) にしてから送信する。このように元の情報を冗長にすることを通信路符号化 (channel encoding) または単に符号化 (encoding) という。符号化された情報にはノイズが混ざって通信路から出力されるが、冗長性をうまく利用し、この出力情報を元の情報に復元する過程を復号化 (decoding) という。

2.2.1 もっとも簡単な符号化とパリティ検査符号

符号化を理解するために、最も簡単な例として同じ情報 (ビット列) を 3 回繰り返して送ることを考える。送られてきた 3 つのビット列が完全に一致していれば、ノイズは無かったものと見なせるし、特定のビットが食い違っていれば、3 つの情報列の中における、そのビット上で多数決を行い、0 か 1 を決めればよい。こうして信号を冗長化することにより、ある程度ノイズが混じっても元の情報を推定することができるようになる。これが最も簡単な方法であり、かつ、伝達路を無駄に使った最も洗練されていない方法である。

これとは別に、もっと洗練された方法としてパリティ検査符号 (parity-check code) というのがある。よくコンピューターのメモリーを購入したときにパリティ付きかパリティなしかが言及されていることがあるが、それはまさにパリティ検査符号がなされているかどうかである。

たとえば，元の信号で7ビットを組にしてその中に1が偶数個か奇数個かに応じて8ビット目に0か1を加えて符号を構成する．そうすると8ビットの組の中に含まれる1の数は必ず偶数個になる．通信路のノイズは0を1に変えるか，1を0に変える．ノイズがあまり強くなって，せいぜい1ビットしか乱さないなら，出力の8ビットの中にはノイズが含まれていないことになるし，奇数個ならビット列のどこかにノイズが入ったことになり，誤りを検出することができる．ただし，誤り検出だけでは誤りを訂正することができないので，誤りが検出された場合は，データを最初から送り直さなくてはならない．誤りの検出された出力データから誤りを訂正したい場合には，送信する前に符号化しさらに受信してから復号化することが必要になる．

本節では，この符号化と復号化について解説する．特に Sourslas 符号を用いた場合における統計力学的な扱いについて紹介する．

2.2.2 2体相互作用の Sourslas 符号

本節では，Sourslas 符号 [39] の基本的な概念を紹介する．統計力学では，扱いを簡単にするために，0と1の代わりに ± 1 を用いることが多い．さらにこの ± 1 を Ising スピンと言うことが多い．0と1の mod 2 での和がビット列演算の基本であるが，これは0を $S_i = 1$ ，そして1を $S_i = -1$ と対応させることにより，Ising スピンの積に対応している．例えば， $0 + 1 = 1$ は Ising スピンにおいて $1 \times (-1) = -1$ に対応しているし， $1 + 1 = 0$ は $(-1) \times (-1) = 1$ になる．このように，ビット列と Ising スピン配位は同一視することができるので，本節においても，0，1ビット列は用いずに，Ising スピンとして取り扱う．

スピンの積で相互作用を作った符号を通信路に送る符号を，Sourslas 符号と呼ぶ． i 番目のスピンと j 番目のスピンの積を J_{ij}^{in} とおけば，

$$J_{ij}^{\text{in}} = \xi_i \xi_j \quad (i \neq j) \quad (2.19)$$

となる．スピンの数を N とし，相互作用の数を N_B とおけば， $N_B > N$ となるので，情報は冗長化されている．情報を式(2.19)のように冗長化した符号を Sourslas 符号と呼ぶ．ただし，この場合は相互作用の数が二つだけの場合である．相互作

用が二つ以上の場合は次の節で説明する．

通信路には $J^{\text{in}} = \{J_{ij}^{\text{in}}\}$ を送信し，出力として $J^{\text{out}} = \{J_{ij}^{\text{out}}\}$ を得る．この J^{out} を復号することで，推定データ $\sigma = \{\sigma_i\}$ を求める．推定データ σ_i は次の状態更新式の平衡状態として求めることができる．

$$\sigma_i(t+1) = \text{sgn} \left(\sum_{j \neq i} J_{ij}^{\text{out}} \sigma_j(t) \right) \quad (2.20)$$

2 体相互作用の Surlas 符号のハミルトニアンは

$$H = - \sum_i \sum_j J_{ij}^{\text{out}} \sigma_i \sigma_j \quad (i \neq j) \quad (2.21)$$

で与えられる．情報が正しいとき，つまり全ての i において $\sigma_i = \xi_i$ であるとき，もしくは全ての i において $\sigma_i = -\xi_i$ であるとき，式 (2.21) が最小値を取ることが分かる．つまり正しい情報が得られたときは，式 (2.21) が基底状態をとるので，正しい情報に復号化する過程は式 (2.21) の緩和過程であることに等しい．

2 体相互作用の Surlas 符号のハミルトニアン (2.21) は，一つのパターンを記憶した場合の連想記憶モデル (2.3 節参照) のハミルトニアンと等しいものになっている．しかし，複数のパターンを記憶した場合の連想記憶モデルのハミルトニアンと，Surlas 符号のハミルトニアンに比べれば，連想記憶モデルの方がより一般化したものとなっている．したがって，歴史的には連想記憶モデル [37] の方が古いモデルであるが，ハミルトニアンに視点を置くことで，連想記憶モデルは Surlas 符号の拡張モデルとなっていると見なせる．

通信路の伝送効率を考慮し，冗長性を下げた場合を考え，最近接相互作用のみで積をとるとする．このようなモデルを Mattis 模型という．Mattis 模型でのハミルトニアンは

$$H = - \sum_{\langle i,j \rangle} J_{ij}^{\text{out}} \sigma_i \sigma_j \quad (2.22)$$

となる． $\langle i,j \rangle$ は最近接相互作用のみで和をとることを意味する．例えば二次元正方形の Mattis 模型の場合，相互作用の数は $N_B = 2N$ となり，やはりスピンの数 N よりも多く情報は冗長化されていることになる．

通信路の出力においては，ところどころ入力符号が反転している．Mattis 模型の場合，元の相互作用で任意の閉じたループ c について J_{ij}^{out} の積を作ると，各

ξ_i が必ず二回ずつ出てくるから積 $f_c = \prod J_{ij}^{\text{out}} = \prod (\xi_i \xi_j)$ は常に 1 である．したがって，雑音の入っていない Mattis 模型ではフラストレーションがないが，雑音による相互作用の符号の反転でフラストレーションが生じる．このよう場合でも，雑音の影響が小さいなら，基底状態が元の正しいスピン配位になっている．このようにスピンの積で作った相互作用を送れば少々雑音があっても，復号化により元のスピン配位が正しく推定できる．

2.2.3 一般の Surlas 符号

前節では 2 体相互作用の Surlas 符号を解説したが，本節では一般の r 体の相互作用を持つ Surlas 符号モデルを紹介する．本節では，Ising スピンのスピン配位 ξ が確率分布 $P(\xi)$ によって生成されるとする． $P(\xi)$ を事前分布と呼ぶ．

r 個のスピンの積

$$J_{i_1 \dots i_r}^{\text{in}} = \xi_{i_1} \cdots \xi_{i_r} \quad (= \pm 1) \quad (2.23)$$

を N_B 個作り，それらを通信路に入力する． $J^{\text{in}} = \{J_{i_1 \dots i_r}^{\text{in}}\}$ に対応する通信路の出力を $J^{\text{out}} = \{J_{i_1 \dots i_r}^{\text{out}}\}$ とする．このような通信路を二元対通信路という．二元通信路の出力特性は，次のような条件付き確率としてあらわせる．

$$P_{\text{out}}(J_{i_1 \dots i_r}^{\text{out}} | \xi_{i_1} \cdots \xi_{i_r}) = \frac{\exp(\beta_p J_{i_1 \dots i_r}^{\text{out}} \xi_{i_1} \cdots \xi_{i_r})}{2 \cosh \beta_p} \quad (2.24)$$

ここで， β_p は次の式で決まるとする．

$$\exp(2\beta_p) = \frac{1-p}{p} \quad (2.25)$$

このように決めることで，式 (2.24) の分母は，次のようになる．

$$\text{式 (2.24) の分母} = 2 \cosh \beta_p \quad (2.26)$$

$$= \exp(\beta_p) + \exp(-\beta_p) \quad (2.27)$$

$$= \sqrt{\exp(2\beta_p)} + \sqrt{\exp(-2\beta_p)} \quad (2.28)$$

$$= \sqrt{\frac{1-p}{p}} + \sqrt{\frac{p}{1-p}} \quad (2.29)$$

$$= \frac{1}{\sqrt{p(1-p)}} \quad (2.30)$$

式 (2.30) を用いれば，入力通りに $J_{i_1 \dots i_r}^{\text{in}} = \xi_{i_1} \cdots \xi_{i_r}$ が出力側においても反転しない確率は，

$$P_{\text{out}}(J_{i_1 \dots i_r}^{\text{out}} | \xi_{i_1} \cdots \xi_{i_r}) = \sqrt{p(1-p)} \exp(\beta_p) \quad (2.31)$$

$$= \sqrt{p(1-p)} \sqrt{\frac{1-p}{p}} \quad (2.32)$$

$$= 1-p \quad (2.33)$$

となる．逆に反転する確率は

$$P_{\text{out}}(J_{i_1 \dots i_r}^{\text{out}} | \xi_{i_1} \cdots \xi_{i_r}) = \sqrt{p(1-p)} \exp(-\beta_p) \quad (2.34)$$

$$= \sqrt{p(1-p)} \sqrt{\frac{p}{1-p}} \quad (2.35)$$

$$= p \quad (2.36)$$

となる．よって，式 (2.24) は通信路の特性をあらわす条件付き確率であることがわかる．

式 (2.24) が各スピンの組み $(i_1 \cdots i_r)$ についてそれぞれ独立に成立するとしよう．つまり，それぞれのビットが互いに独立に雑音の影響を受け，さらに $1 \rightarrow -1$ に反転する確率と $-1 \rightarrow 1$ に反転する確率が同じである通信路を考える．このような通信路を無記憶通信路 (memoryless binary symmetric channel) と呼ぶ．このとき，送信する情報全体としての雑音は式 (2.24) の積になる．

$$P_{\text{out}}(J^{\text{out}} | \xi) = \frac{1}{(2 \cosh \beta_p)^{N_B}} \exp(\beta_p \sum_{i_1 \dots i_r} J_{i_1 \dots i_r}^{\text{out}} \xi_{i_1} \cdots \xi_{i_r}) \quad (2.37)$$

和は式 (2.24) でスピンの積を作った組 $(i_1 \cdots i_r)$ 全てについてとる． N_B はこの和に現れる項の数であり，通信路に入力されるビット数 N に等しい．

2.2.4 最大事後確率法とハミルトニアン

出力 $J_{i_1 \dots i_r}^{\text{out}}$ から元の信号 ξ の推定 σ をできるだけ正しく求めることを考える．このためには出力 $J_{i_1 \dots i_r}^{\text{out}}$ が与えられたときの推定値 σ についての条件付き確率 (事後確率) $P(\sigma | J^{\text{out}})$ が必要になる．

事前分布を $P_{\text{in}}(\sigma)$ として、ベイズの公式を使うと、

$$P(\sigma|J^{\text{out}}) = \frac{P_{\text{out}}(J^{\text{out}}|\sigma)P_{\text{in}}(\sigma)}{\text{Tr}_{\sigma}P_{\text{out}}(J^{\text{out}}|\sigma)P_{\text{in}}(\sigma)} \quad (2.38)$$

となる。 $P(\sigma|J^{\text{out}})$ は、式 (2.37) であらわされることが既にわかっている。事前分布 $P_{\text{in}}(\sigma)$ が分かれば、式 (2.38) による元の信号の推定が実行できる。一般的な取り扱いを可能にするために、さまざまな情報が同じ確率で生成されているような情報源を考える。このとき、 $P_{\text{in}}(\sigma)$ は定数となるので、

$$P(\sigma|J^{\text{out}}) = \frac{\exp(\beta_p \sum^{N_B} J_{i_1 \dots i_r}^{\text{out}} \sigma_{i_1} \dots \sigma_{i_r})}{\text{Tr}_{\sigma} \exp(\beta_p \sum^{N_B} J_{i_1 \dots i_r}^{\text{out}} \sigma_{i_1} \dots \sigma_{i_r})} \quad (2.39)$$

となる。

σ を求めるにおいて一番もってもらいたいものは、式 (2.39) を最大化するような σ である。分母は定数となるので、式 (2.39) を最大化するためには分子を最大化すればよい。つまり、分子の \exp の括弧内を最大化すればよい。分子の \exp の括弧内は、統計物理学で Boltzmann 因子と呼ばれる。Boltzmann 因子を最大化することは、系の基底状態を求めることに相当する。

よって、ハミルトニアンを

$$H = - \sum^{N_B} J_{i_1 \dots i_r}^{\text{out}} \sigma_{i_1} \dots \sigma_{i_r} \quad (2.40)$$

とおくことで、系の基底状態を求めることとハミルトニアンを最小化することは等価になる。つまり一番もってもらいたい σ を求めることとハミルトニアン (2.40) を最小化することは同じことになり、同時にこれは基底状態探索問題に帰着されたことになる。このような方法を最大事後確率 (MAP) 法と呼ぶ。

修復値は、 σ_i の式 (2.39) による期待値

$$\langle \sigma_i \rangle = \text{Tr}_{\sigma} \sigma_i P(\sigma|J^{\text{out}}) \quad (2.41)$$

に符号関数をかけたもの

$$\text{sgn}(\langle \sigma_i \rangle) = \text{sgn}(\text{Tr}_{\sigma} \sigma_i P(\sigma|J^{\text{out}})) \quad (2.42)$$

とすればよい。

2.2.5 有限温度復号

MAP はビット列のベクトル σ 全体として事後確率を最大化するが、これとは異なる復号化法もある。特定のビット i に注目し、その他のビットは当面考慮しないことにする。式 (2.39) の確率分布で σ_i 以外のスピンについて和を取ると、 σ_i だけについての事後確率が得られる。

$$P(\sigma_i | J^{\text{out}}) = \frac{\text{Tr}_{\sigma(\neq \sigma_i)} \exp(\beta_p \sum J_{i_1 \dots i_r}^{\text{out}} \sigma_{i_1} \dots \sigma_{i_r})}{\text{Tr}_{\sigma} \exp(\beta_p \sum J_{i_1 \dots i_r}^{\text{out}} \sigma_{i_1} \dots \sigma_{i_r})} \quad (2.43)$$

式 (2.43) において $P(\sigma_i = 1 | J^{\text{out}})$ と $P(\sigma_i = -1 | J^{\text{out}})$ を比べ、前者が大きいなら i 番目のビットの復号結果を 1 とし、後者が大きいなら、 -1 とする。このプロセスを全てのビットについて行って、得られた結果を集めて最終的な復号結果とする。このような方法を有限温度復号 (finite-temperature decoding) あるいは MPM(Maximizer of Posterior Marginals) という。

修復値は、 σ_i の式 (2.43) による期待値

$$\langle \sigma_i \rangle_{\beta_p} = \sum_{\sigma=\pm 1} \sigma_i P(\sigma_i | J^{\text{out}}) \quad (2.44)$$

に符号関数をかけたもの

$$\text{sgn}(\langle \sigma_i \rangle_{\beta_p}) = \text{sgn} \left(\sum_{\sigma=\pm 1} \sigma_i P(\sigma_i | J^{\text{out}}) \right) = \text{sgn} \left(\frac{\text{Tr}_{\sigma} \sigma_i P(\sigma_i | J^{\text{out}})}{\text{Tr}_{\sigma} P(\sigma_i | J^{\text{out}})} \right) \quad (2.45)$$

とすればよい。

ビット列全体の事後確率の最大化が MAP であるのに対し、各ビットごとの事後確率の最大化が有限温度復号である。

2.2.6 本節のまとめ

本節では、Sourlas 符号による符号訂正を情報統計力学的な観点から解説した。特に 2 体相互作用の場合は、冗長化された送信信号 J^{in} が、一つのパターンを記憶した場合の連想記憶モデルの学習則と同一のものとなる。実際に、2 体相互作用の Sourlas 符号と一つのパターンを記憶した場合の連想記憶モデルのハミルトニアンは等しいものになることを次節において紹介する。

実際に Surlas は Hopfield 模型の影響を多分に受けており，歴史的には連想記憶モデルの方が古いが，連想記憶モデルの方が Surlas 符号の拡張モデルなっていると云える．連想記憶モデルについては次の 2.3 節において解説する．

2.3. 連想記憶モデル

人をはじめ，さまざまな生体システムは，現在のコンピュータが及びも抜かない実に巧みで柔軟な情報処理機能を持つ．その基本となる構成単位は，ニューロン (neuron, 神経細胞) と呼ばれ，多数の樹状突起が出ており，その中の一本は長くのびて先端がいくつかの枝に別れ，他のニューロンの樹状突起につながっている．この接合部をシナプスと呼んでいる．ニューロンは，さまざまな状況に応じて情報を電気信号 (活動電位) としてシナプスを通して他のニューロンへ伝える．このとき，ニューロンの内部電位があるしきい値を超えないと，発火して他のニューロンに信号が伝わらない仕組みになっている．また，他の多数のニューロンからの信号が一つのニューロンに伝わることで電位が決まる．

生体システムは以上のようなニューロンが多数結合して，一つの情報処理システムを構成している．これをニューラルネットワーク (neural network, 神経回路網) と呼ぶ．パターン認識の問題が，実はパーセプトロン (perceptron) というモデルによって解決できるとして，F.Rosenblatt が 1950 年代後半に発表している [25]．特に，正しい類別が行われるところまで，学習が収束するという理論的裏づけもあって，多くの研究がなされた．ところが，M.Minsky と S.Papert が，パーセプトロンは線形分離可能なパターンに限って正しく類別が出来て識別が可能なのであって，非線形が入った場合もはや不可能であることを理論的に示した．これによってニューラルネットワークの研究も下火になったが，1986 年 D.E.Rumelhart ら [38] が誤差逆伝播法 (Back-Propagation, 4.7 節参照) という新しい学習則を提案したことにより，再びニューラルネットワークの研究に火がついた．

また，ニューラルネットワークのもう一つの典型モデルとして，連想記憶モデル [26]-[29],[32] がある．これは，ある記憶事象の一部だけが与えられていた場合に，それをキーとして全体を想起するような心的過程を神経回路モデルを用いてモデル化したものである．近年，この連想記憶モデルに統計物理学の解析法が有効的であることが分かり [33], [37]，多くの統計物理学者がニューラルネットワークの研究に従事した．本節では，この連想記憶モデルについて紹介する．

2.2 節では 2 体相互作用の Sourlas 符号と連想記憶モデルの統計力学的取り扱い

いにおける関連性について述べたが，先にも述べたとおり連想記憶モデルの方が Sourlas 符号よりも歴史的には古く，Sourlas は Hopfield 型の連想記憶モデル [26] の影響を多分に受けている．本節では，実際に 2 体相互作用の Sourlas 符号のハミルトニアンと一つのパターンを記憶した場合の連想記憶モデルのハミルトニアンが等しいものになっていることを見ていくとともに，連想記憶モデルの性質を実際のシミュレーションを用いることで紹介する．

2.3.1 相互結合型ネットワーク

ニューラルネットワーク (以下単にネットワークと略す) の系を構成する N 個の素子 (ニューロン) のそれぞれが，相互に結合しているネットワークを相互結合型ネットワークと呼ぶ．相互結合型ネットワークでは素子の出力が入力にフィードバックされるため，そのフィードバックの回数に対応して時刻 t が自然に導入される．時刻 t におけるネットワークの状態を状態ベクトル $S(t)$ によりあらわす．このベクトルの次元はネットワークが持つ素子の総数 N に等しく，その要素 $S_j, (1 \leq j \leq N)$ は各素子の状態をあらわしている．ある時刻における状態はその時刻における出力になる．この様子を概略的に図に示すとすれば，図 2.1 のようになる．

各素子は他の素子から送られてくる信号の合計値を，入力 (内部状態または膜電位) として受け入れ，さらにその値を出力関数に通して出力値とする．通常，出力関数には次式で定義される符号関数 $\text{sgn}(x)$ が用いられる．

$$\text{sgn}(x) = \begin{cases} +1 & (x \geq 0) \\ -1 & (x < 0) \end{cases} \quad (2.46)$$

また，符号関数をグラフであらわすと，図 2.2 のようになる．

状態ベクトルの各要素 S_i が常に ± 1 をとるので，出力関数にも ± 1 をとる符号関数を用いている．実際の神経細胞との対応を考えると，素子の状態 ± 1 に応じて素子の出力が $S_i(t) = +1$ である時は素子が発火の状態にあり， $S_i(t) = -1$ である時は非発火の状態にあると解釈される．

素子 j と素子 i を結ぶ結合荷重を J_{ij} とすれば，状態 $S_i(t)$ を $S_i(t+1)$ に更新す

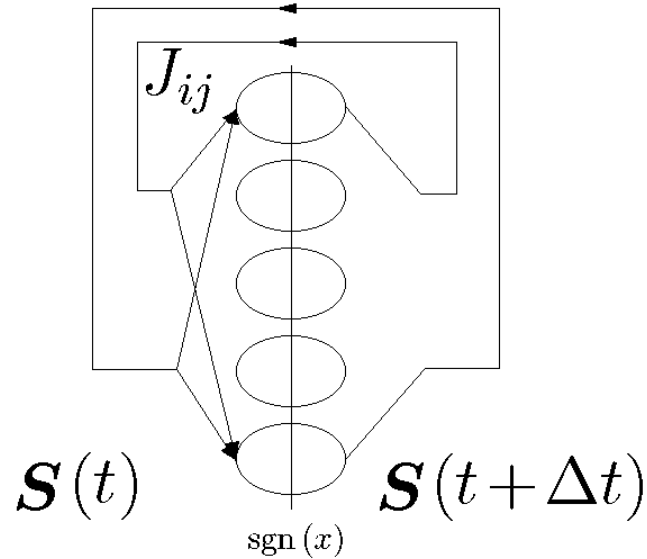


図 2.1 相互結合型ネットワークの概略図

る状態方程式を次の式により定式化することができる．

$$S_i(t+1) = \text{sgn} \left(\sum_{j \neq i} J_{ij} S_j(t) \right) \quad (2.47)$$

素子の出力は離散的な時刻 t で同期的に変化する．式 (2.47) の状態更新のダイナミクスを離散時間同期型ダイナミクスと呼ぶ．この漸化式によって状態を次々に更新していくと，系の状態はやがて平衡状態や，リミットサイクルと呼ばれる周期状態に落ち着く．

2.3.2 ヘップ則による学習

ある記憶事象の一部だけが与えられており，これをキーとして全体を想起するような過程を，連想記憶と呼ぶ．我々の経験から考えると，このような過程が脳の中でも行われている可能性がある．ここではその心的過程を状態更新式 (2.47) を用いてモデル化することを考える．

実際に連想記憶を実現するには，式 (2.47) における結合荷重 J_{ij} をうまく決めなくてはならない．式 (2.47) が連想記憶モデルなるように， ξ^μ により結合荷重

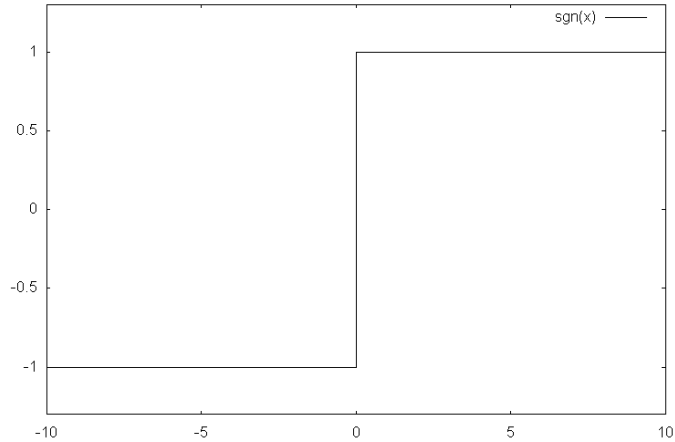


図 2.2 符号関数

J_{ij} を決める最も有効な方法の一つとしてヘップ則がある．ヘップ則は次の式によりあらわされる．

$$J_{ij} = \frac{1}{N} \sum_{\mu=1}^P \xi_i^{\mu} \xi_j^{\mu} \quad (i \neq j) \quad (2.48)$$

ここで， P はモデルに記憶させる記憶パターンの数である．自己結合 J_{ii} は考えないものとし， 0 と置くことにする．記憶パターン数 P は，モデルを構成する素子の数 N に対する比ではかると，モデルの性能を議論する上で便利な場合がある．そこでその比

$$\alpha = \frac{P}{N} \quad (2.49)$$

を記憶率とする．式 (2.47) で示される自己想起型の構造を持ち，結合荷重 J_{ij} を式 (2.48) の相関学習で決めたモデルを連想記憶モデルと呼ぶ．

記憶パターンの候補としては，画像のビットパターンやそのパターンをある程度特徴抽出したパターンなどが考えられる．記憶パターンは各要素が $+1$ または -1 をとる N 次元ベクトルであるが，理論的な扱いを簡単にするために，記憶パターンをランダムに生成することにする．すなわち， μ 番目の記憶パターン ξ^{μ} の i 番目の成分 ξ_i^{μ} を次の確率で決める．

$$\text{Prob}[\xi_i^{\mu} = \pm 1] = \frac{1}{2} \quad (2.50)$$

記憶パターンを式 (2.50) で生成されるようなランダムパターンとすることに関しては批判があり、この批判にはもっともな点もある。しかしながら、たとえば記憶パターンに画像のビットパターンを用いてその解析をしても、その結果がモデルの性質によるものなのか、記憶パターンの性質によるものかはっきりしない。ランダムパターンによる解析を行っておけば、少なくとも記憶パターンの数学的な性質が単純になるので、解析結果が理解しやすい。また、記憶パターンにランダムパターンを用いた解析によって得られた知見は、画像ビットパターンを用いたモデルに対しても有効である。

2.3.3 記憶パターンの想起

ここでは連想記憶モデルが記憶パターンを想起できる理由を定性的な説明を行う。1番目の記憶パターン ξ^1 を想起する場合を考える。想起する ξ^1 を想起パターンと呼ぶ。モデルの時刻 t での状態 $S(t)$ と記憶パターンである 1番目の記憶パターン ξ^1 とのオーバーラップ $m(t)$ を次のように定義する。

$$m(t) = \frac{1}{N} \sum_{i=1}^N \xi_i^1 S_i(t) \quad (2.51)$$

記憶パターンとモデルの状態ベクトルはともに各成分が ± 1 の N 次元ベクトルであるから、オーバーラップ $m(t)$ はそれらの間の方向余弦である。オーバーラップのように素子に関する平均操作で求められる変数を巨視的変数と呼ぶ。 M , ($0 < M \leq 1$) を用いて、初期状態 $S(t)$ を以下の確率でランダムに決める。

$$\text{Prob}[S_i(0) = \pm 1] = \frac{1 \pm M \xi_i^1}{2} \quad (2.52)$$

このようにして生成された初期状態 $S(0)$ と記憶パターン ξ^1 のオーバーラップ $m(0)$ は、次のようにして与えられる。まず、記憶パターンは式 (2.50) を用いて独自に生成される。よって、想起パターンの各成分 ξ_i^1 は素子の番号 i に関して独立である。また式 (2.52) により、初期状態 $S(0)$ の各成分 $S_i(0)$ も素子の番号 i に関して独立である。 $N \rightarrow \infty$ の極限を考えているので、大数の法則から初期オーバ

ラップ $m(0)$ は M に収束する .

$$m(0) = \lim_N \frac{1}{N} \sum_{i=1}^N \xi_i^1 S_i(0) = E[\xi_i^1 S_i(0)] = M \quad (2.53)$$

$$1 < m(0) \leq 1 \quad (2.54)$$

時刻 $t = 0$ の状態 $S(0)$ から決まる時刻 $t = 0$ の i 番目の素子の内部状態 $u_i(0)$ は ,

$$u_i(0) = \sum_{j \neq i}^N J_{ij} S_j(0) \quad (2.55)$$

$$= \xi_i^1 m(0) + \frac{1}{N} \sum_{\mu=2}^{\alpha N} \sum_{j \neq i}^N \xi_i^\mu \xi_j^\mu S_j(0) \quad (2.56)$$

で与えられる . 式 (2.56) の初項は後で述べる理由から想起のためのシグナルと考えられる . 第 2 項は想起パターン以外の記憶パターンからの干渉をあらわし , クロストークノイズと呼ばれる . このように内部状態を想起パターンからの寄与のシグナル項と , それ以外の記憶パターンからの寄与のクロストークノイズ項に分ける解析法は , シグナルノイズ解析 (S/N 解析) と呼ばれている . これは物理学で多体問題を議論する際に用いられる一体近似の考え方と同じである . 特に , $N \rightarrow \infty$ の極限では , 式 (2.56) の第 2 項のクロストークノイズ項は平均 0 分散 α のガウス分布に従う (先ほど導入した記憶率 α がクロストークノイズをあらわすパラメータとして現れたことに注意せよ) . したがって記憶させたパターンの数が小さい (α が小さい) とき , 式 (2.56) の第 2 項を 0 と近似できるので ,

$$S_i(1) = \text{sgn}(u_i(1)) = \text{sgn}(\xi_i^1 m(0)) = \xi_i^1 \quad (2.57)$$

となり , 時刻 $t = 1$ で 1 番目の記憶パターンの想起が完了する . この解析からわかるように , 式 (2.55) の初項は想起のためのシグナルであり , 時刻 $t = 0$ においてガウス分布で記述されるクロストークノイズが想起の妨害をすることがわかる . もし時刻 $t = 1$ で完全に想起されなくても , 式 (2.47) に従って繰り返し計算すれば , $m(0) < m(1) < \dots < m(\infty) = 1$ のように想起できる可能性がある . 以上が , 連想記憶モデルが記憶パターンの想起能力を持つ定性的な理由である .

2.3.4 連想記憶モデルのハミルトニアン

Hopfield は、式 (2.46) において非同期に更新される場合のハミルトニアンが

$$H = -\frac{1}{2} \sum_i \sum_j J_{ij} S_i S_j = -\frac{1}{2} \sum_i S_i \sum_j J_{ij} S_j \quad (2.58)$$

で与えられることを示した [37]。ここで、 S_i は i 番目の素子の出力をあらわす。素子の状態は、時刻 t において全素子が同時に更新されるのではなく、各時刻に一つの素子の状態のみ更新される。

非同期の連想記憶モデルの状態更新は式 (2.58) の緩和過程になっている。つまり、ネットワークは初期条件（入力パターン）に最も近いハミルトニアンの極小状態に落ち着いて変化が停止する。これを証明ために、 i 番目の素子 S_i が更新されて、 S'_i になる場合を考える。このとき、 S'_i は式 (2.46) より、

$$S'_i = \text{sgn}\left(\sum_j J_{ij} S_j\right) \quad (2.59)$$

と与えられる。

$S'_i = S_i$ の場合は明らかにハミルトニアンは変化しない。そこで $S'_i = -S_i$ と更新された場合のハミルトニアン H' と更新前のハミルトニアン H の差を考える。

$$H' - H = -\sum_{j \neq i} J_{ij} S'_i S_j + \sum_{j \neq i} J_{ij} S_i S_j \quad (2.60)$$

$$= 2S_i \sum_{j \neq i} J_{ij} S_j \quad (2.61)$$

$$= 2S_i \sum_j J_{ij} S_j - 2J_{ii} \quad (2.62)$$

ここで、式 (2.59) より、式 (2.62) の第一項が負になることが分かる。また、 P を記憶パターン数、 N を素子数とすれば、ヘップ則 (2.48) より $J_{ii} = P/N$ であるので、第二項も負になることが分かる。よって素子が更新されるたびに、ハミルトニアンが減少していくことが分かる。

このように式 (2.48) と式 (2.58) を組み合わせ、非同期の状態が更新される連想記憶モデルを、Hopfield の功績にちなんで Hopfield モデルと呼ぶ。Hopfield の最大の功績は、式 (2.62) のハミルトニアンを導入することで連想記憶モデルのダイ

ナミクスの議論を可能にしたことである．Sourlas はこの Hopfield のこの仕事に触発され，Sourlas 符号のモデルを開発した．実際に記憶パターンが一つの場合，ハミルトニアン の式 (2.58) は

$$H = -\frac{1}{2} \sum_i \sum_j J_{ij} S_i S_j = -\frac{1}{2} \sum_i \sum_j \xi_i \xi_j S_i S_j \quad (2.63)$$

となり，2 対相互作用の Sourlas 符号のハミルトニアン の式 (2.21) と一致する．Hopfield の仕事は Sourlas 符号だけではなく，情報科学の広い分野まで影響を与えた．2.1 節や 2.2 節でも述べたように，画像修復問題や Sourlas 符号もハミルトニアンを定義することで，問題を基底状態探索問題に置き換えることが可能になった．このように，ハミルトニアンを定義することで定性的な解析をする方法を統計力学的手法と呼ぶことが多い．

2.3.5 連想記憶モデルの性質とシミュレーション

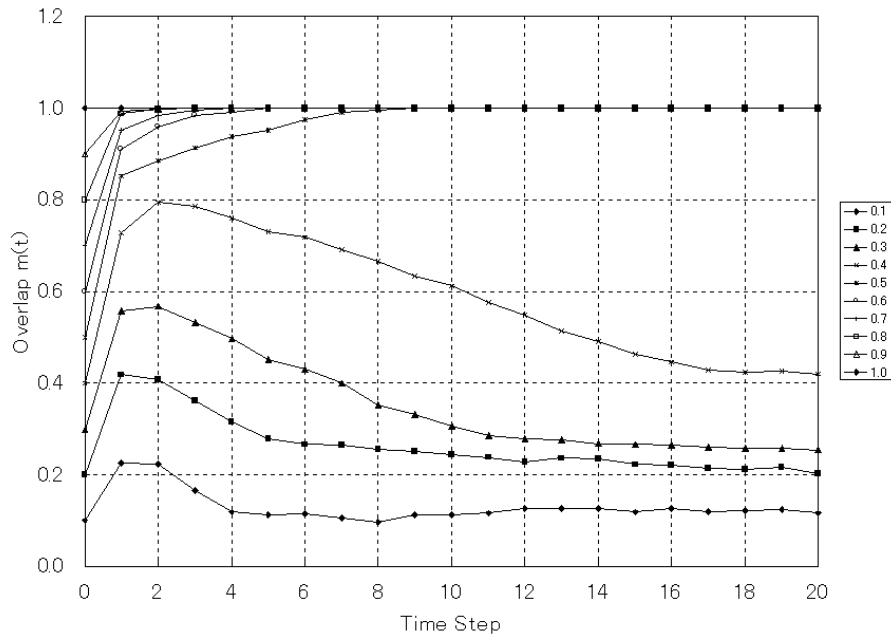


図 2.3 素子数 $N = 1000$ ，パターン数 $P = 120$ ， $(\alpha = 0.12)$ のときの想起の様子．横軸が時刻 t をあらわし，縦軸が各時刻でのオーバーラップ $m(t)$ をあらわす．

本節では、実際に連想記憶モデルのシミュレーションを行うことでその定性的な性質を見ていく。素子数 $N = 1000$ ，パターン数 $P = 120$ ($\alpha = 0.12$) のときの想起の様子を図 2.3 に示す。横軸が時刻 t をあらわし，縦軸が各時刻でのオーバーラップ $m(t)$ をあらわす。初期オーバーラップ $m(0)$ が 1 に近い場合は時刻 $t = 1$ において $m(1) = 1$ となり，想起に成功していることがわかる。また， $m(0) \geq 0.5$ の場合は，時刻 $t = 1$ において想起が完了していなくても，式 (2.47) にしたがって繰り返し計算すれば想起に成功することが分かる。

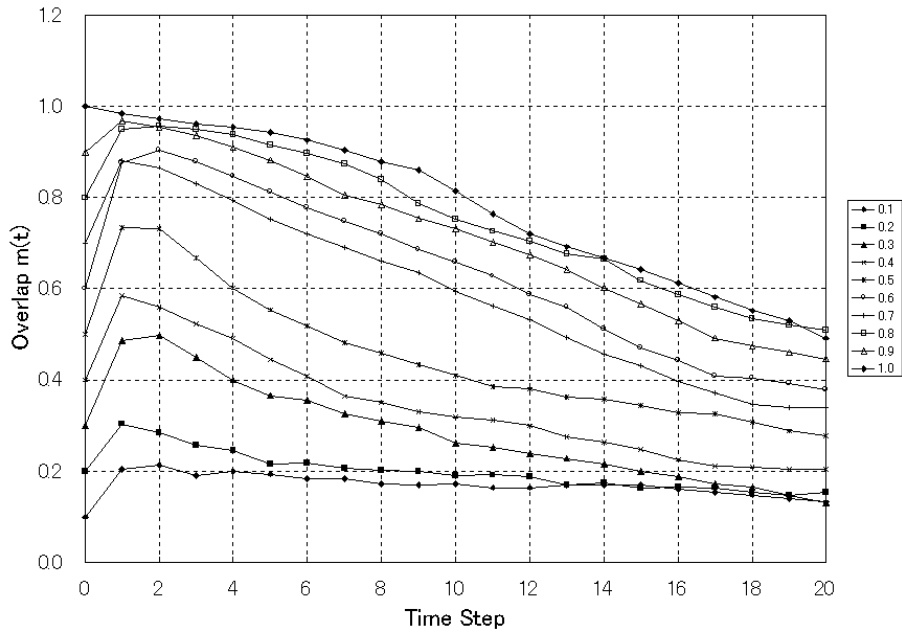


図 2.4 素子数 $N = 1000$ ，パターン数 $P = 200$ ，($\alpha = 0.20$) のときの想起の様子。縦軸が時刻 t をあらわし，横軸が各時刻でのオーバーラップ $m(t)$ をあらわす。

次に，素子数 $N = 1000$ ，パターン数 $P = 200$ ($\alpha = 0.20$) のときの想起の様子を図 2.4 に示す。記憶パターン自身を初期状態にしても ($m(0) = 1$)，モデルの状態は記憶パターンから離れていく。これらの結果から記憶パターン数 P が 120 から 200 に増える間に，モデルの性質が変化していることがわかる。

この様子をもう少し詳しく見るために，記憶率 α を横軸にとり，系の平衡状態におけるオーバーラップ ($m(\infty)$) を縦軸にとったグラフを図 2.5 に示す。ただし，初期状態は $m(0) = 1$ とした。

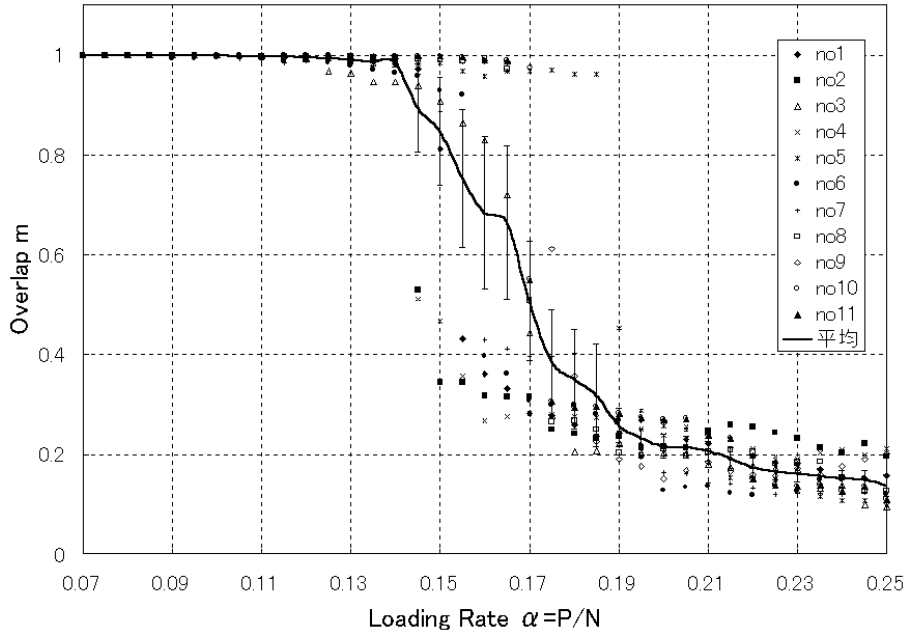


図 2.5 連想記憶モデルの相転移をあらわすグラフ．横軸は記憶率 α ，縦軸は系が平衡状態に達したときのオーバーラップ $m(\infty)$ をあらわす．初期状態を $m(0) = 1$ とした場合において平衡状態に達したときのオーバーラップ $m(\infty)$ を見た．実線は 11 回のシミュレーションにより得られたデータの平均をあらわし，エラーバーは標準偏差をあらわす．

このグラフから記憶率 α に関して，臨界的な記憶率 $\alpha_c \approx 0.14$ が存在することがわかる．その α_c の前後でモデルの性質が急激に変化している．このような変化を我々は相転移 [42] と呼んでいる．記憶率が $\alpha < \alpha_c$ である場合，想起パターン ξ^1 は式 (2.47) のダイナミクスの安定的な平衡状態であり，その平衡状態は有限の大きさの引き込み領域をもつ．一方 $\alpha > \alpha_c$ では，想起パターン ξ^1 は不安定化し，初期状態を想起パターンに設定しても系の状態 $S(t)$ は想起パターンから離れていく．この記憶率 α_c を記憶容量と呼ぶ．式 (2.47), (2.48) で記述される連想記憶モデルが想起を成功させるには，記憶率が $\alpha < \alpha_c$ である必要がある．また，ここまでの議論では一つの想起パターン ξ^1 について議論しているが，全ての記憶パターン ξ^μ は式 (2.48) の相関学習に関して対称である．つまり記憶率 α が記憶容量 α_c を超えると全ての記憶パターンが不安定化する．

偽記憶と引き込み領域

図 2.3 から初期状態でのオーバーラップが $m(0) \leq 0.4$ の場合、想起に失敗していることがわかる。想起に失敗したときのモデルの挙動を調べると、想起に成功した場合と同様に平衡状態に達している。この平衡状態は、記憶パターンの1つであることもあるし、そうでないこともある。記憶パターンに含まれない平衡状態を偽記憶と呼ぶ。しかし本来、モデルが記憶パターン自体を覚えていないので、図 2.3 のようにオーバーラップを使って記憶パターンを想起したか否かの判断はできない。

図 2.5 で記憶率が記憶容量より大きい場合 ($\alpha > \alpha_c$) にパターンの想起がうまく行かなくなるのは、連想記憶モデルが記憶パターン以外に非常に多くの偽記憶を持つことに由来している。偽記憶は、素子数 N に対し $O(\exp N)$ 個存在することが知られている [35]。

また、図 2.3 から初期オーバーラップ $m(0) \geq 0.5$ であるとき想起に成功し、 $m(0) \leq 0.4$ では想起に失敗していることがわかる。つまり、 $0.4 < m(0) < 0.5$ の間に記憶パターンに引き込まれるか否かの境界が存在する。この臨界的な初期オーバーラップ m_c を臨界オーバーラップと呼ぶ。初期状態のオーバーラップが臨界オーバーラップよりも大きく、最終的に記憶パターンに引き込まれる初期状態の領域を引き込み領域と呼ぶ。引き込み領域は入力である初期状態 $S(0)$ に含まれる誤りを訂正する能力をあらわす。

ところで、このように偽記憶へ収束することは連想記憶モデルの問題の一つになっている。第 4 章では、ハイブリッド型の学習を用いることで、記憶パターンにより大きな引き込み領域を持たせ、偽記憶に収束しにくいモデルを提案する。

高速状態更新式

実際のシミュレーションを行うに当たりその速度は重要な問題である。そこでここでは、状態更新を高速に遂行することが可能な高速状態更新式 [27] について解説する。

式 (2.47) を高速状態更新式に変換するにあたり、まず最初に状態の表記法を変

換する．式 (2.47) の表記法では，発火状態にある素子と非発火状態にある素子の状態値をそれぞれ $S_i = \pm 1$ とあらわしてきたが，ここではそれらの値をそれぞれ $\sigma_i = 1, 0$ と表記を変える．それに応じて記憶パターンの状態値 $\xi_i^\mu = \pm 1$ も $\zeta_i^\mu = 1, 0$ と置き換える．

c 言語において符号なし整数型を宣言した場合，現在ほとんどのコンピュータが 1 ワードにおいて 32 ビットを確保している．つまり，この表記法の変換により 1 ワードに 32 個の素子の状態値を記憶できるのである．したがって，コンピュータのメモリーを大幅に節約することが可能になると同時に，次に説明する論理演算を用いた計算も可能になり処理速度を高めることができる．

素子の状態の表記法を換えたことにより，式 (2.47) と式 (2.48) を次の 1 つの式に置き換えることができる．

$$\sigma_i(t+1) = \frac{2}{NP} \sum_{\mu=1}^P \sum_{j=1}^N \zeta_i^\mu \oplus \zeta_j^\mu \oplus \sigma_j(t) \quad (2.64)$$

ここで NP により最後に行われる割り算は整数型であり，小数点以下はすべて切り捨てられる．また印 \oplus は XOR 論理演算（排他的論理和）を示している．この XOR 論理演算は並列処理を可能にする．つまり 1 ワード（32 ビット）に記録されている 32 個のニューロンが式 (2.64) により一度に更新される．同様に， μ 番目の記憶パターン ζ_j^μ と時刻 t における状態のオーバーラップ $M_\mu(t)$ を次のように表現することができる．

$$M_\mu(t) = \sum_{j=1}^N \zeta_j^\mu \oplus \sigma_j(t) \quad (2.65)$$

N が有限の場合

これまでは $N \rightarrow \infty$ の場合における連想記憶モデルの性質について考えてきた．特に，図 2.3-2.5 のシミュレーションでは $N \rightarrow \infty$ の近似として $N = 1000$ を用いた．これは $N = 1000$ で確認される現象が $N \rightarrow \infty$ においても成り立つという経験的な常識に基づいたものである．

ここでは $N \ll \infty$ の場合におけるモデルの性質について考えていく．実際に素子数が少ない場合の動作を確認するために，まずは $N = 320$ としてシミュレー

ションを行う。 P 個のパターンを記憶させ、様々な初期オーバーラップ $m(0)$ から初めて目的のパターンに収束する成功率を測定する。

図 2.3, 図 2.4 のように N が大きい場合は一つの系が統計的な振る舞いを示すのに対し, $N \ll \infty$ の場合はゆらぎが大きく一つの系に統計的な振る舞いをとらせるのは難しい。したがって, それぞれのシミュレーション設定において系に偏りが生じることになるので, これを回避するためになるべく多くの回数のシミュレーションを行いその統計結果を求める。ここでは 1000 回のシミュレーションを行い, その結果の平均を統計データとして用いることにする。シミュレーションには高速状態更新式 (2.64) を用いた。

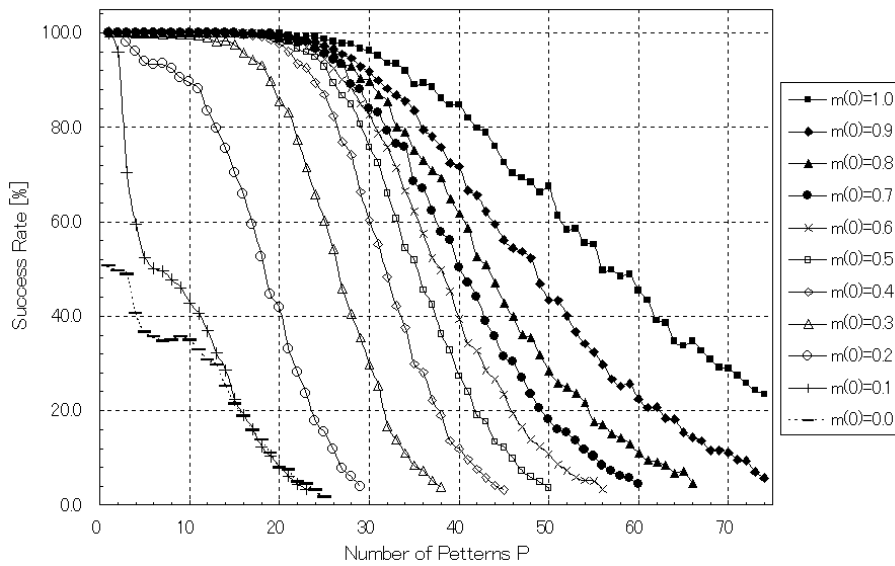


図 2.6 素子数 $N = 320$, 様々な初期オーバーラップ $m(0)$ から開始した場合における想起の成功率。横軸が記憶パターン数 P をあらわし, 縦軸が想起の成功率をあらわす。

様々な初期オーバーラップ $m(0)$ から開始した場合における想起の成功率を図 2.6 に示す。縦軸が想起の成功率をあらわし, 横軸が記憶パターン数 P をあらわす。図 2.6 から, 初期オーバーラップが高いほど想起の成功率が高いことが確認できる。また, 記憶パターン数 P が少ないほど (記憶率 α が小さいほど) 成功率が高くなっていることが確認できる。

次に， N の変化に伴って性質がどのように変化するかを見る．記憶パターン自身を初期状態 ($m(0) = 1$) とした場合の想起の成功率を図 2.7 に示す．横軸を記憶率 α とし，縦軸を想起の成功率としている．1000 回のシミュレーション結果の平均を用いている．それぞれの折れ線は N を 160 から 1920 まで増やしていった場合の想起の成功率をあらわしている．

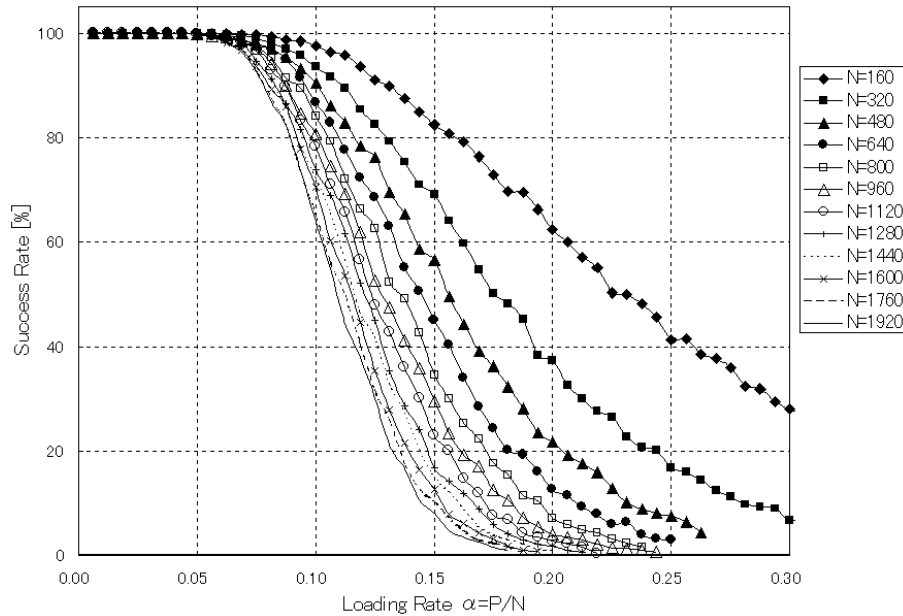


図 2.7 記憶パターン自身を初期状態にして ($m(0) = 1$) 想起の成功率をあらわすグラフ．縦軸が想起の成功率，横軸が記憶率 α をあらわす．

素子数 N が少ない場合 ($N = 160$) は P が大きくなってても想起の成功率が下がっていない．それに対し， N が大きい場合 ($N = 1920$) は， $\alpha = 0.14$ 付近において，ほとんど想起が成功していない．この事実から N が小さいほど記憶容量 α_c が大きくなることがわかる．

連続型の相互結合型ネットワーク

ここまでのモデルでは素子の状態値として離散値 ± 1 を持つ離散型モデルをとり扱ってきたが，ここからは素子の状態値が連続値を持つ連続型モデルについて

て解説する．連続値を持つ素子の方が離散値を持つ素子よりも，現実的な神経細胞により近いといえる．また，電気回路などへの応用が容易で，場合によっては解析が簡単になる [32] ．

連続型モデルの場合は離散型モデルと異なり，内部状態と出力状態を明確に区別する．後述する状態方程式には内部状態が用いられていることから，モデルの系は主に内部状態により議論される．時刻 t におけるネットワークの内部状態を内部状態ベクトル $x(t)$ によりあらわす．このベクトルの次元はネットワークが持つ素子の総数 N に等しく，その要素 $x_j, (1 \leq j \leq N)$ は各素子の内部状態をあらわしている．離散型モデルの場合とは異なり，この内部状態は，原則としていかなる実数をもとりうる．

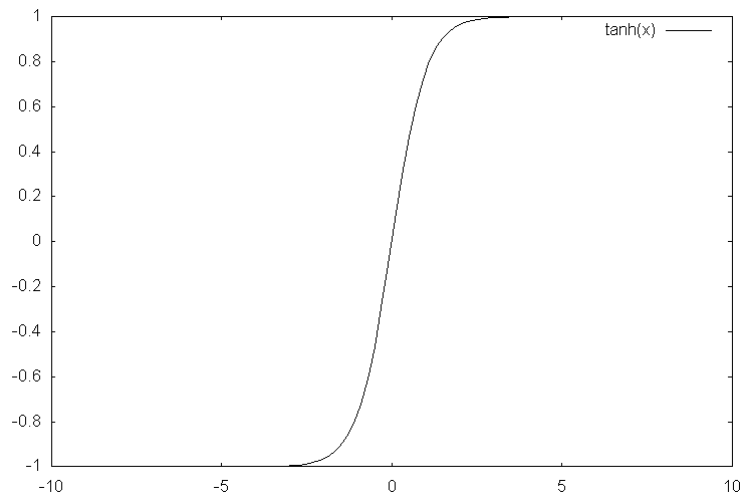


図 2.8 双曲正接

$g(s)$ を出力関数ベクトルとすれば，出力状態は $g(x(t))$ とあらわされる．内部状態ベクトルと同様に出力関数ベクトルの次元は N に等しく，ベクトルの各要素 $g_i(s_i)$ は各素子の出力関数をあらわす．出力関数には通常非線形な飽和関数が用いられるが，出力関数として双曲正接 $g_i(x) = \tanh(a_i x)$ を用いる．ここで， a_i は正の定数である．双曲正接をグラフであらわすと図 2.8 のようになる．出力関数に双曲正接を用いることにより出力状態のとりうる値は $-1 < x_i(t) < +1$ となる．

連続値を持つモデルの状態方程式は次のような微分方程式によりあらわされる．

$$\frac{dx_i}{dt} = -\frac{1}{\tau}x_i + \sum_{j=1}^m J_{ij}g_i(x_j) + b_i, \quad (i = 1, \dots, N) \quad (2.66)$$

J_{ij} は離散型モデルの場合と同様， j 番目から i 番目の素子への結合の重みをあらわし，連続型モデルにおいてもヘップ則 (2.48) を用いる． τ は時定数， b_i はバイアスである．式 (2.66) の τ を左辺に移項して，

$$\tau \frac{dx_i}{dt} = -x_i + \sum_{j=1}^m J_{ij}g_i(x_j) + b_i, \quad (i = 1, \dots, N) \quad (2.67)$$

のようにあらわすこともある．式 (2.67) のようにあらわすことの利点は，系が平衡状態の場合，すなわち式 (2.67) の左辺が 0 になった場合において，内部状態 $x_i(0)$ が τ に依存しない点である．実際，第 4 章において隠れ素子を導入する際に式 (2.67) を用いる方が都合がよいので，第 4 章ではモデルの状態方程式としては式 (2.67) を用いている．

双曲正接の傾き a_i が $a_i \rightarrow \infty$ であるとき，式 (2.66) は離散型モデルの状態更新式 (2.47) と等価になることが一般に知られている [26]．

2.3.6 本節のまとめ

本節では，Sourlas 符号の拡張と見なすことのできる連想記憶モデルについての解説をした．Sourlas 符号と同様に Hopfield 型の連想記憶モデルのハミルトニアンを定義することで想起過程がハミルトニアンの緩和過程に相当することを示した．特に，2 体相互作用の Sourlas 符号のハミルトニアンと一つのパターンを記憶した場合の連想記憶モデルのハミルトニアンが等しいものになっていることを示した．

さらにシミュレーションを用いることで連想記憶モデルの主な性質を見てきた．その中で偽記憶への収束が連想記憶モデルの問題の一つになっていることを紹介した．この偽記憶への収束を緩和させるモデルの一つとして，第 4 章ではハイブリッド型連想記憶モデルについて提案する．

第3章 画像修復の統計力学

3.1. この章のはじめに

符号化することで元の情報を修復することが可能であるが、既に 2.1 節で述べたように、実際には符号化できない場合が存在する。本章では、符号化しない場合の情報復元問題として再び画像修復問題を取り扱う。そして 2.1 節で述べた画像修復問題をさらに一般化し、統計力学的な議論を行うことで画像問題の定性的な性質を見ていく。

画像処理の歴史は長く、コンピュータのめざましい発達に伴いその技術は急速に高まりつつある。しかも、画像処理技術の需要はコンピュータに留まらず、携帯電話、デジタルカメラ、ビデオカメラに始まり、冷蔵庫、洗濯機、電子レンジ等の生活に身近な家電製品への実装の可能性も議論されるに至っている。このような状況において、従来の個々の用途に突起した形での緻密な画像処理技術だけでは対応しきれない状況が増加し、むしろ多少の緻密さを犠牲にしてもより柔軟で頑健な画像処理技術に対する社会的要請とともにベイズ統計を用いた画像処理への期待が高まりつつある。

一般に画像処理技術は、一見、全く関係のないように見える統計力学と様々な点で共通点が見いだされる。このことを見るためにまず従来の典型的画像処理技術について概観する。画像は正方格子のような規則格子上に並んだ画素という基本単位から構成され、各画素に割り当てられた階調値 (デジタル画像では $0, 1, 2, \dots, 255$ の整数値であらわされる。) をもとにディスプレイ上の光の強さが決められる。デジタル画像処理においては画像に対する修復、エッジ検出、領域分割、圧縮等の目的に応じてフィルターというものが設計される。画像修復を目的として用いられる基本的フィルターとして平滑化フィルターとメジアンフィ

ルターがある．いずれのフィルターも各画素ごとにその近傍画素の集合を考える．劣化画像における各画素ごとにその近傍画素について階調値の平均値を計算し，これを修復画像の対応する画素の階調値として採用するのが平滑化フィルターである．この平均値の代わりにその近傍画素の集合について階調値のメジアンを修復画像の対応する画素の階調値として採用するのがメジアンフィルターである．これらはいずれも「近傍画素の状態からもとの画像の注目画素の階調値を推定しよう」という方針に基づいており，さらに修復画像は「各画素の階調値は近傍画素の階調値と著しく違うことはない」ように構成される．

一方，統計力学において磁性材料の研究に用いられてきた確率モデルについて概観する．磁性体の物性を説明する確率モデルは規則格子上に並んだスピンというなれば原子単位の小さな磁石を基本単位として構成される．磁性体全体がどのような磁石になるかはこのスピンという基本単位がそれぞれどのような状態にあるかによって決まる．最も基本的な確率モデルにおいてはスピンは，極が上を向いているか下を向いているかの二つの状態のみであらわされる．更に，スピンごとに近傍スピンの定義され，各スピンは対応する近傍スピンと相関があり，近傍スピンの状態によってどちら向きを向くかが決定される．すなわち「各スピンの状態は近傍スピンと同じ状態を向く傾向があるように全体として配列する」わけである．この二つの状況を並べると以下の対応が成立する．

画素 \leftrightarrow スピン

近傍画素 \leftrightarrow 近傍スピン

このことから磁性体の物性を説明するものとして用いられてきた確率モデルを画像処理に応用できないものかと考えることはごく自然なことである．実際，このような試みは新しい画像処理技術としてコンピュータサイエンスに従事する研究者により 1985 年ごろから様々の試験研究が行われてきた [1]-[3]．基本公式としてベイズの公式が用いられていることから理論的基盤を統計学者が支えてきたわけである [4]-[6],[11]-[14]．1995 年以降，ようやく物理学者からもそのモデル構造の共通性から研究対象として注目されるに至っている [7][10]．

しかしながら，2.1 節で述べたようなこれまでの研究は，画像に重畳されるノイズが画素毎に独立であると仮定されており [11]-[24]，空間的な相関を持つノイ

ズを重畳するようなモデルはまだ存在しない．しかし，光学系の特性等を考慮すれば，重畳されるノイズに空間的な相関が生ずると考えるのは自然である．そこで本章では，空間的に並進対称な相関を持つノイズ下での画像修復を統計力学的な観点から議論する．具体的には，原画像とノイズが並進対象性を持つような多重ガウス分布に従って生成される場合を議論する．この場合，フーリエ変換を用いることで，フーリエ空間における多重ガウス分布を対角化することができる．

次に，空間的な相関を持たないノイズモデルが，相関のあるモデルにどれだけ対応できるかを議論する．上で述べたように，これまでの研究が相関を持たないノイズで画像修復を行っていた背景には，ハイパーパラメータであるノイズの分散等を適切に選べば，ほどほどに良い画像修復が行われるのではないかという希望的な予測があったからである．ここでは，その予測がどの程度妥当であるかを理論的に議論する．ハイパーパラメータの推定方法としては，以下の2つの方法を考える．

一つ目の方法は，周辺尤度最大化によるハイパーパラメータの推定を行う方法である [17],[18]．この方法は，劣化画像から求まる周辺尤度と呼ばれる量を最大化するようにハイパーパラメータを決定する方法で，基本的には劣化画像さえ分かれば実現可能な方法である．二つ目の方法は，原画像と修復画像の二乗誤差を最小にするようなハイパーパラメータを求める方法である．この方法は，原画像が分かっている必要がないので，現実的には実行不可能な方法である．

ノイズの生成確率と，修復で仮定したノイズが同じモデルである場合，これら二つの方法は同じ結果を与えることが知られている．しかし，今回のように修復の際のモデルが生成モデルと異なる場合，二つの結果が一致しないことを示す．

3.2. モデル

3.2.1 画像の生成過程

画像は通常2次元格子上で定義されるが，ここでは一般化して d 次元格子上の画像を考える． d 次元立方格子の一边あたりの画素数を L とする．この場合，素子の総数 N は $N = L^d$ となる．

格子点 i 上の原画像の画素値を ξ_i とおく．本章では，確率分布にしたがって画像が生成されると仮定する．特に，以下の多重ガウス分布にしたがって生成されるとする．

$$P(\boldsymbol{\xi}) = \frac{1}{Z_{\text{prior}}(\beta, h)} \exp[-H(\boldsymbol{\xi}, \beta, h)] \quad (3.1)$$

$$H(\boldsymbol{\xi}, \beta, h) = \boldsymbol{\xi}^T(\beta G + hI)\boldsymbol{\xi} \quad (3.2)$$

ただし， β, h は正のスカラーである．式 (3.1) の左辺を事前分布と呼ぶことが多い． G は行列， I は単位行列であり， $Z_{\text{prior}}(\beta, h)$ は

$$Z_{\text{prior}}(\beta, h) = (2\pi)^{\frac{N}{2}} |2(\beta G + hI)|^{-\frac{1}{2}} \quad (3.3)$$

である．式 (3.2) の H は統計力学でいうハミルトニアンである．

本章では行列 G が並進対称行列であるとし，その要素 $G_{i,j}$ を

$$G_{i,j} = 2d\delta_{(i-j),0} - \sum_{\boldsymbol{\delta}} \delta_{(i-j),-\boldsymbol{\delta}} - \sum_{\boldsymbol{\delta}} \delta_{(i-j),\boldsymbol{\delta}} \quad (3.4)$$

とした場合を議論する [16]．ただし， d 次元ベクトル $\boldsymbol{\delta}$ は次の値をとる．

$$\boldsymbol{\delta} = (1, 0, \dots, 0), (0, 1, 0, \dots, 0), \dots, (0, \dots, 0, 1) \quad (3.5)$$

一方で式 (3.4) の $\delta_{i,j}$ はクロネッカーのデルタであり，次で定義される．

$$\delta_{i,j} = \begin{cases} 1 & (i = j) \\ 0 & (i \neq j) \end{cases} \quad (3.6)$$

式 (3.4) から $G_{i,j}$ は生成される原画像に最近接画素間のみの相互作用をもたらすことを意味する．したがって，式 (3.1) の確率で生成される原画像では， β が大きな値をとるときは第 1 項の最近接相互作用の効果が大きくなるので，隣り合った画素が同じ値をとろうとする．一方， h は画素値の振幅をあらわし， h が大きいほど画素値の絶対値は小さくなる傾向がある．

以後フーリエ変換が取り扱えるように，本章で取り扱う画像にはすべて周期境界条件が成り立つと仮定する．式 (3.5) を用いれば周期境界条件を次のように表現できる．

$$\xi_i = \xi_{i+L\boldsymbol{\delta}} \quad (3.7)$$

以上が画像の生成過程の解説である．以下，画像の劣化過程について解説する．

3.2.2 画像の劣化過程

$\tau = \{\tau_i\}$ を劣化画像する．原画像と同様に， d 次元格子上的画像であるとし， d 次元立方格子の一辺あたりの画素数を L とする．素子の総数 N も原画像と同様に $N = L^d$ となる．劣化過程も多重ガウス分布に従うとする．よって，ガウス雑音による劣化過程を次のように置く．

$$P_{\text{out}}(\tau|\xi) = \frac{1}{Z_{\text{noise}}} \exp \left[-\frac{1}{2} (\tau - \xi)^T R^{-1} (\tau - \xi) \right] \quad (3.8)$$

ただし，

$$Z_{\text{noise}}(R) = (2\pi)^{\frac{N}{2}} |R|^{\frac{1}{2}} \quad (3.9)$$

である．また， $(\tau - \xi)$ は原画像に重畳されるノイズ画像である．式 (3.8) の左辺を尤度と呼ぶことが多い．

R はノイズの相関行列であり，並進対称性を持つとする．今までの研究では空間的には相関を持たないノイズが用いられており [11]-[24]，これは，

$$R_{i,j} = v^2 \delta_{i,j} \quad (3.10)$$

に対応する． v^2 がノイズの分散に相当する．

本章では式 (3.10) に，ピクセル間の距離の二乗 $(i - j)^2$ に対して \exp で減少するような項を加えることで，次のような相関行列 R を議論する．

$$R_{i,j} = (1 - a)b^2 \delta_{i,j} + ab^2 \exp \left[-\frac{(i - j)^2}{\kappa^2} \right] \quad (3.11)$$

ただし， $0 \leq a \leq 1$ である．

式 (3.11) において $a = 0$ と置けば第二項が 0 となり，式 (3.10) で表現される空間的な相関を持たないノイズと等価になる．また， $i = j$ のとき， a の値とは無関係に $R_{i,i} = b^2$ となる．すなわち， a が 1 に近づくほど，ノイズにおける各ピクセル間の相関性が強まり，ノイズのシミの大きさが大きくなる．逆に a が 0 に近づくほど，ノイズにおける各ピクセル間の相関性が弱まり，ノイズのシミの大きさは小さくなっていく (3.6 節，3.8.1 項参照)．

3.3. ベイズ推定と画像修復の有効ハミルトニアン

ここで、原画像 ξ と劣化画像 τ が与えられた場合、ベイズの公式により次が成り立つ。

$$P(\xi|\tau) = \frac{P_{\text{out}}(\tau|\xi)P(\xi)}{\int d\sigma P_{\text{out}}(\tau|\xi)P(\xi)} \quad (3.12)$$

ベイズ推定では $P(\xi|\tau)$ を事後分布、 $P_{\text{out}}(\tau|\xi)$ を尤度、 $P(\xi)$ を事前分布と呼ぶことがある。

σ を修復画像とする。式 (3.12) において原画像 ξ と修復画像 σ を入れ換えても構わないので、この入れ換えにより劣化画像 τ が与えられた場合の修復画像 σ の事後確率が計算できる。

$$P(\sigma|\tau) = \frac{P_{\text{out}}(\tau|\sigma)P(\sigma)}{\int d\sigma P_{\text{out}}(\tau|\sigma)P(\sigma)} \quad (3.13)$$

$$= \frac{\exp[-H_{\text{eff}}]}{\int d\sigma \exp[-H_{\text{eff}}]} \quad (3.14)$$

ただし、

$$H_{\text{eff}} = \sigma^T(\beta G + hI)\sigma + \frac{1}{2}(\tau - \sigma)^T R^{-1}(\tau - \sigma) \quad (3.15)$$

であり、統計力学では有効ハミルトニアンと呼ばれる。事後確率 $P(\sigma|\tau)$ を最大化するには、式 (3.13) の右辺の分子を最大化すれば良く、このためには有効ハミルトニアン (3.15) を最小化すれば良いことになる。つまり、よりもっともらしい修復画像 σ を求めることは有効ハミルトニアン (3.15) の基底状態探索問題に帰着される。このような方法を最大事後確率 (Maximum A Posteriori, MAP) 法と呼ぶ。

さらに、画像修復における式 (3.12) の意味を考える。 $P_{\text{out}}(\tau|\xi)$ は原画像 ξ から劣化画像 τ が生成される確率である。 $P(\xi)$ は原画像 ξ そのものの出現確率であり、ベイズ統計においては事前確率と呼ばれる。ベイズの公式 (3.12) の右辺は、事前確率 $P(\xi)$ をもとに、まず原画像 ξ が生成され、その原画像 ξ から劣化過程 $P_{\text{out}}(\tau|\xi)$ を通して劣化画像 τ が生成されるといういわゆる順過程の状況をあらわしている。観測された劣化画像はこの順過程によって生成されたものであると考えれば、ベイズの公式 (3.12) は「順過程をあらわす式」が「劣化画像 τ が与えられたという条件のもとでの原画像 ξ に対する確率 $P(\xi|\tau)$ 」に等しいということ

を意味している．すなわち，順過程に対する仮定が適切なものであればそれだけ良好な推定を行うことができるということになる．もちろん，劣化画像が生成された過程を徹底的に分析し，その分析結果をすべて順過程の過程に反映させることができればきわめて良好な推定が可能となる．しかしその場合，分析を行ったシステムに対しては極めて緻密な推定を行うことができるが，そのシステムから逸脱した状況においては良好な推定を行うことが困難となってしまう．用途がある特定の場合作に限定されているならそれで良いのであるが，不特定多数の顧客を相手にした柔軟で頑健な知的情報処理システムとして構成しようとする場合，順過程における仮定を最小限にとどめながらしかもその仮定が顧客の要求の本質をついたものであることが必要となる．

3.4. 理論

3.4.1 修復画像の期待値

式 (3.4),(3.11) で与えられる相関行列はいずれも並進対称行列であり，これらはフーリエ変換により対角化できる．本章では，この対角化を利用して計算を簡単化することで理論的な解析を行う．

離散フーリエ変換を次で定義する．

$$\tilde{\xi}_{\mathbf{k}} = \frac{1}{\sqrt{N}} \sum_{\mathbf{j}} \xi_{\mathbf{j}} e^{-i\mathbf{k} \cdot \mathbf{j}} \quad (3.16)$$

逆フーリエ変換 (フーリエ表記) を次で定義する．

$$\xi_{\mathbf{j}} = \frac{1}{\sqrt{N}} \sum_{\mathbf{k}} \tilde{\xi}_{\mathbf{k}} e^{i\mathbf{k} \cdot \mathbf{j}} \quad (3.17)$$

i は虚数単位である． \mathbf{k} も \mathbf{j} と同じ d 次元ベクトルであり，自由度は L^d である．

$$\sum_{\mathbf{j}} = L^d = N, \quad \sum_{\mathbf{k}} = L^d = N \quad (3.18)$$

また， \mathbf{k} の各成分は以下のような値をとる．

$$0, \frac{2}{L}\pi, \frac{4}{L}\pi, \dots, \frac{2(L-1)}{L}\pi \quad (3.19)$$

式 (3.16) を用いると , 式 (3.1) の exp 内をフーリエ表記することができて

$$P(\boldsymbol{\xi}) = \frac{1}{Z_{\text{prior}}(\beta, h)} \exp \left[- \sum_{\mathbf{k}} (\beta \tilde{G}_{\mathbf{k}} + h) \tilde{\xi}_{\mathbf{k}} \tilde{\xi}_{-\mathbf{k}} \right] \quad (3.20)$$

$$\tilde{G}_{\mathbf{k}} = \sum_{\delta} [2 - 2 \cos(\mathbf{k} \cdot \delta)] \quad (3.21)$$

となる .

式 (3.11) の $R_{i,j}$ も並進対称行列であるので , $G_{i,j}$ と同様にフーリエ変換を用いた対角化ができる .

$$\tilde{R}_{\mathbf{k}} = (1 - a)b^2 + ab^2 \sum_l e^{-\frac{l^2}{\kappa^2}} \cos(\mathbf{k} \cdot \mathbf{l}) \quad (3.22)$$

ただし , ベクトル l の要素 l がとりうる範囲は ,

$$-(L - 1) \leq l \leq L - 1 \quad (3.23)$$

とする . 式 (3.8) の exp 内も式 (3.1) と同様にフーリエ表記できるので ,

$$P_{\text{out}}(\boldsymbol{\tau}|\boldsymbol{\xi}) = \frac{1}{Z_{\text{noise}}} \exp \left[-\frac{1}{2} \sum_{\mathbf{k}} \tilde{R}_{\mathbf{k}}^{-1} (\tilde{\tau}_{\mathbf{k}} - \tilde{\xi}_{\mathbf{k}})(\tilde{\tau}_{-\mathbf{k}} - \tilde{\xi}_{-\mathbf{k}}) \right] \quad (3.24)$$

となる . このようにフーリエ変換を用いることで行列を対角化することができるので以降の計算が容易になる .

式 (3.14) の事後確率 $P(\boldsymbol{\sigma}|\boldsymbol{\tau})$ に基づく期待値を $\langle \cdot \rangle$ であらわす . 修復画素 σ_j の期待値は ,

$$\langle \sigma_j \rangle = \frac{1}{\sqrt{N}} \sum_{\mathbf{k}} \langle \tilde{\sigma}_{\mathbf{k}} \rangle e^{i\mathbf{k} \cdot \mathbf{j}} \quad (3.25)$$

となる . フーリエ成分 $\sigma_{\mathbf{k}}$ の期待値は , 式 (3.14) を用いて ,

$$\langle \tilde{\sigma}_{\mathbf{k}} \rangle = \int \prod_{\mathbf{k}'} d\tilde{\sigma}_{\mathbf{k}'} \tilde{\sigma}_{\mathbf{k}} P(\boldsymbol{\sigma}|\boldsymbol{\tau}) \quad (3.26)$$

$$= \frac{\int \prod_{\mathbf{k}'} d\tilde{\sigma}_{\mathbf{k}'} \tilde{\sigma}_{\mathbf{k}} e^{-\hat{H}_{eff}}}{\int \prod_{\mathbf{k}'} d\tilde{\sigma}_{\mathbf{k}'} e^{-\hat{H}_{eff}}} \quad (3.27)$$

$$= \frac{1}{\int d\tilde{\sigma}_{\mathbf{k}} e^{-\hat{H}_{eff}}} \int d\tilde{\sigma}_{\mathbf{k}} \tilde{\sigma}_{\mathbf{k}} e^{-\hat{H}_{eff}} \quad (3.28)$$

$$\begin{aligned}
&= \frac{1}{\int d\tilde{\sigma}_{\mathbf{k}} e^{-\hat{H}_{eff}}} \int d\tilde{\sigma}_{\mathbf{k}} \tilde{\sigma}_{\mathbf{k}} \\
&\quad \times \exp \left[-(\hat{\beta}\tilde{G}_{\mathbf{k}} + \hat{h})\tilde{\sigma}_{\mathbf{k}}\tilde{\sigma}_{-\mathbf{k}} \right. \\
&\quad \left. - \frac{1}{2\hat{R}_{\mathbf{k}}}(\tilde{\tau}_{\mathbf{k}} - \tilde{\sigma}_{\mathbf{k}})(\tilde{\tau}_{-\mathbf{k}} - \tilde{\sigma}_{-\mathbf{k}}) \right] \quad (3.29)
\end{aligned}$$

となる．ただし， \hat{H}_{eff} は式 (3.15) をフーリエ表記したもので，

$$\hat{H}_{eff} = \sum_{\mathbf{k}} (\hat{\beta}\tilde{G}_{\mathbf{k}} + \hat{h})\tilde{\sigma}_{\mathbf{k}}\tilde{\sigma}_{-\mathbf{k}} + \sum_{\mathbf{k}} \frac{1}{2\hat{R}_{\mathbf{k}}}(\tilde{\tau}_{\mathbf{k}} - \tilde{\sigma}_{\mathbf{k}})(\tilde{\tau}_{-\mathbf{k}} - \tilde{\sigma}_{-\mathbf{k}}) \quad (3.30)$$

である．また，ここでは画像やノイズの生成確率をあらわす $\beta, h, \tilde{R}_{\mathbf{k}}$ が既知ではないとしている．ただし，それらの関数形は既知であるとし，それらに対応するハイパーパラメータを $\hat{\beta}, \hat{h}, \hat{R}_{\mathbf{k}}$ としている．特に， $\hat{R}_{\mathbf{k}}$ を k によらないとすれば，式 (3.10) の相関を持たないノイズに帰着する．

ここで，

$$\hat{A}_{\mathbf{k}} = \hat{\beta}\tilde{G}_{\mathbf{k}} + \hat{h} + \frac{1}{2\hat{R}_{\mathbf{k}}}, \quad (3.31)$$

$$\hat{B}_{\mathbf{k}} = \frac{1}{2\hat{R}_{\mathbf{k}}} \quad (3.32)$$

を式 (3.29),(3.30) に代入すれば，

$$\begin{aligned}
&\langle \tilde{\sigma}_{\mathbf{k}} \rangle \\
&= \frac{\int d\tilde{\sigma}_{\mathbf{k}} \tilde{\sigma}_{\mathbf{k}} \exp \left[-\hat{A}_{\mathbf{k}} \left| \tilde{\sigma}_{\mathbf{k}} - \frac{\hat{B}_{\mathbf{k}}}{\hat{A}_{\mathbf{k}}} \tilde{\tau}_{\mathbf{k}} \right|^2 \right]}{\int d\tilde{\sigma}_{\mathbf{k}} \exp \left[-\hat{A}_{\mathbf{k}} \left| \tilde{\sigma}_{\mathbf{k}} - \frac{\hat{B}_{\mathbf{k}}}{\hat{A}_{\mathbf{k}}} \tilde{\tau}_{\mathbf{k}} \right|^2 \right]} \quad (3.33)
\end{aligned}$$

$$= \frac{\hat{B}_{\mathbf{k}}}{\hat{A}_{\mathbf{k}}} \tilde{\tau}_{\mathbf{k}}, \quad (3.34)$$

となり，式 (3.25) の $\langle \sigma_j \rangle$ は，

$$\langle \sigma_j \rangle$$

$$= \frac{1}{\sqrt{N}} \sum_{\mathbf{k}} \frac{\hat{B}\mathbf{k}}{\hat{A}\mathbf{k}} \tilde{\tau}_{\mathbf{k}} e^{i\mathbf{k}\cdot\mathbf{j}} \quad (3.35)$$

$$= \frac{1}{\sqrt{N}} \sum_{\mathbf{k}} \frac{\frac{1}{2\hat{R}\mathbf{k}}}{\hat{\beta}\tilde{G}\mathbf{k} + \hat{h} + \frac{1}{2\hat{R}\mathbf{k}}} \tilde{\tau}_{\mathbf{k}} e^{i\mathbf{k}\cdot\mathbf{j}} \quad (3.36)$$

$$= \frac{1}{N} \sum_{\mathbf{k}} \sum_{\mathbf{i}} \frac{\tau_{\mathbf{i}} \frac{1}{2\hat{R}\mathbf{k}} e^{i\mathbf{k}\cdot(\mathbf{j}-\mathbf{i})}}{\hat{\beta}\tilde{G}\mathbf{k} + \hat{h} + \frac{1}{2\hat{R}\mathbf{k}}} \quad (3.37)$$

$$= \frac{1}{N} \sum_{\mathbf{k}} \sum_{\mathbf{i}} \frac{\tau_{\mathbf{i}} \frac{1}{2\hat{R}\mathbf{k}} \cos[\mathbf{k}\cdot(\mathbf{j}-\mathbf{i})]}{\hat{\beta}\tilde{G}\mathbf{k} + \hat{h} + \frac{1}{2\hat{R}\mathbf{k}}}, \quad (3.38)$$

となる．本章では式 (3.38) を修復画像として用いる．

3.4.2 修復誤差最小化基準

ここでは前項で求めた修復画像と原画像の修復誤差最小化基準で， $\hat{\beta}$, \hat{h} , $\hat{R}_{\mathbf{k}}$ などのハイパーパラメータを推定する．ピクセルごとの修復画像と原画像の差の平均二乗誤差 E_1 の期待値は，

$$E_1 = \left\langle \left\langle \frac{1}{N} \sum_{\mathbf{j}} (\xi_{\mathbf{j}} - \langle \sigma_{\mathbf{j}} \rangle)^2 \right\rangle \right\rangle \quad (3.39)$$

となる．ここで， $\langle \langle \cdot \rangle \rangle$ は同時分布 $P(\tau, \xi) = P_{out}(\tau|\xi)P_s(\xi)$ に関する平均である．式 (3.39) をフーリエ表記すると次のようになる．

$$E_1 = \frac{1}{N} \sum_{\mathbf{k}} \langle \langle (\tilde{\xi}_{\mathbf{k}} - \langle \tilde{\sigma}_{\mathbf{k}} \rangle)(\tilde{\xi}_{-\mathbf{k}} - \langle \tilde{\sigma}_{-\mathbf{k}} \rangle) \rangle \rangle \quad (3.40)$$

$P(\tau, \xi)$ はフーリエ表記では対角化されているので，各々の \mathbf{k} について容易に計算することができる．

$$\begin{aligned} & \langle \langle (\tilde{\xi}_{\mathbf{k}} - \langle \tilde{\sigma}_{\mathbf{k}} \rangle)(\tilde{\xi}_{-\mathbf{k}} - \langle \tilde{\sigma}_{-\mathbf{k}} \rangle) \rangle \rangle \\ &= \frac{1}{Z} \int \int d\tilde{\xi}_{\mathbf{k}} d\tilde{\tau}_{\mathbf{k}} \left| \tilde{\xi}_{\mathbf{k}} - \frac{\hat{B}\mathbf{k}}{\hat{A}\mathbf{k}} \tilde{\tau}_{\mathbf{k}} \right|^2 \end{aligned} \quad (3.41)$$

$$\times \exp \left[-A_{\mathbf{k}} \left| \tilde{\xi}_{\mathbf{k}} - \frac{\tilde{B}_{\mathbf{k}}}{\tilde{A}_{\mathbf{k}}} \tilde{\tau}_{\mathbf{k}} \right|^2 - \left(\tilde{B}_{\mathbf{k}} - \frac{\tilde{B}_{\mathbf{k}}^2}{\tilde{A}_{\mathbf{k}}} \right) |\tilde{\tau}_{\mathbf{k}}|^2 \right] \quad (3.42)$$

$$\begin{aligned} &= \frac{1}{Z} \int \int d\tilde{\xi}_{\mathbf{k}} d\tilde{\tau}_{\mathbf{k}} \left| \tilde{\xi}_{\mathbf{k}} + \tilde{\tau}_{\mathbf{k}} \left(\frac{\tilde{B}_{\mathbf{k}}}{\tilde{A}_{\mathbf{k}}} - \frac{\hat{B}_{\mathbf{k}}}{\hat{A}_{\mathbf{k}}} \right) \right|^2 \\ &\quad \times \exp \left[-\tilde{A}_{\mathbf{k}} |\tilde{\xi}_{\mathbf{k}}|^2 \right] \\ &\quad \times \exp \left[- \left(\tilde{B}_{\mathbf{k}} - \frac{\tilde{B}_{\mathbf{k}}^2}{\tilde{A}_{\mathbf{k}}} \right) |\tilde{\tau}_{\mathbf{k}}|^2 \right] \end{aligned} \quad (3.43)$$

$$\begin{aligned} &= \frac{1}{Z} \int \int d\tilde{\xi}_{\mathbf{k}} d\tilde{\tau}_{\mathbf{k}} \left[|\tilde{\xi}_{\mathbf{k}}|^2 + \left(\frac{\tilde{B}_{\mathbf{k}}}{\tilde{A}_{\mathbf{k}}} + \frac{\hat{B}_{\mathbf{k}}}{\hat{A}_{\mathbf{k}}} \right)^2 |\tilde{\tau}_{\mathbf{k}}|^2 \right] \\ &\quad \times \exp \left[-\tilde{A}_{\mathbf{k}} |\tilde{\xi}_{\mathbf{k}}|^2 \right] \\ &\quad \times \exp \left[- \left(\tilde{B}_{\mathbf{k}} - \frac{\tilde{B}_{\mathbf{k}}^2}{\tilde{A}_{\mathbf{k}}} \right) |\tilde{\tau}_{\mathbf{k}}|^2 \right] \end{aligned} \quad (3.44)$$

$$= \frac{1}{2\tilde{A}_{\mathbf{k}}} + \left(\frac{\tilde{B}_{\mathbf{k}}}{\tilde{A}_{\mathbf{k}}} - \frac{\hat{B}_{\mathbf{k}}}{\hat{A}_{\mathbf{k}}} \right)^2 \frac{1}{2\tilde{B}_{\mathbf{k}} - \frac{2\tilde{B}_{\mathbf{k}}^2}{\tilde{A}_{\mathbf{k}}}} \quad (3.45)$$

ただし,

$$\tilde{A}_{\mathbf{k}} = \beta \tilde{G}_{\mathbf{k}} + h + \frac{1}{2\tilde{R}_{\mathbf{k}}} \quad (3.46)$$

$$\tilde{B}_{\mathbf{k}} = \frac{1}{2\tilde{R}_{\mathbf{k}}} \quad (3.47)$$

$$\begin{aligned} Z &= \int \int d\tilde{\xi}_{\mathbf{k}} d\tilde{\tau}_{\mathbf{k}} \\ &\quad \times \exp \left[-\tilde{A}_{\mathbf{k}} \left| \tilde{\xi}_{\mathbf{k}} - \frac{\tilde{B}_{\mathbf{k}}}{\tilde{A}_{\mathbf{k}}} \tilde{\tau}_{\mathbf{k}} \right|^2 - \left(\tilde{B}_{\mathbf{k}} - \frac{\tilde{B}_{\mathbf{k}}^2}{\tilde{A}_{\mathbf{k}}} \right) |\tilde{\tau}_{\mathbf{k}}|^2 \right] \end{aligned} \quad (3.48)$$

である。したがって,

$$E_1 = \frac{1}{2N} \sum_{\mathbf{k}} \left[\frac{1}{\tilde{A}_{\mathbf{k}}} + \left(\frac{\tilde{B}_{\mathbf{k}}}{\tilde{A}_{\mathbf{k}}} - \frac{\hat{B}_{\mathbf{k}}}{\hat{A}_{\mathbf{k}}} \right)^2 \frac{1}{\tilde{B}_{\mathbf{k}} - \frac{\tilde{B}_{\mathbf{k}}^2}{\tilde{A}_{\mathbf{k}}}} \right] \quad (3.49)$$

$$\begin{aligned}
&= \frac{1}{2N} \sum_{\mathbf{k}} \left[\frac{1}{\beta \tilde{G}_{\mathbf{k}} + h + \frac{1}{2\tilde{R}_{\mathbf{k}}}} \right. \\
&\quad \left. + \left(\frac{\frac{1}{2\tilde{R}_{\mathbf{k}}}}{\beta \tilde{G}_{\mathbf{k}} + h + \frac{1}{2\tilde{R}_{\mathbf{k}}}} - \frac{\frac{1}{2\hat{R}_{\mathbf{k}}}}{\hat{\beta} \tilde{G}_{\mathbf{k}} + \hat{h} + \frac{1}{2\hat{R}_{\mathbf{k}}}} \right)^2 \right. \\
&\quad \left. \times \frac{\beta \tilde{G}_{\mathbf{k}} + h + \frac{1}{2\tilde{R}_{\mathbf{k}}}}{\frac{1}{2\tilde{R}_{\mathbf{k}}} (\beta \tilde{G}_{\mathbf{k}} + h)} \right] \quad (3.50)
\end{aligned}$$

となる．

$$\hat{\beta} = \beta, \hat{h} = h, \hat{R}_{\mathbf{k}} = \tilde{R}_{\mathbf{k}}, \quad (3.51)$$

であるとき，最小二乗誤差 E_1 は最低値，

$$E_{1_{min}} = \frac{1}{2N} \sum_{\mathbf{k}} \left[\frac{1}{\beta G_{\mathbf{k}} + h + \frac{1}{2\tilde{R}_{\mathbf{k}}}} \right] \quad (3.52)$$

をとる．より正確には， $\hat{\beta}, \hat{h}, \hat{R}_{\mathbf{k}}$ と $\beta, h, \tilde{R}_{\mathbf{k}}$ のそれぞれの比が正しければ E_1 は同じ最低値，式 (3.52) をとる．この式 (3.52) は修復の限界をあらわしている．

同様に劣化画像と原画像の平均二乗誤差 E_2 を求めることができる．

$$E_2 = \left\langle \left\langle \frac{1}{N} \sum_j (\tilde{\xi}_j - \tilde{\tau}_j)^2 \right\rangle \right\rangle \quad (3.53)$$

$$= \frac{1}{N} \sum_{\mathbf{k}} \tilde{R}_{\mathbf{k}} = \frac{1}{N} \sum_i \tilde{R}_{i,i} = b^2 \quad (3.54)$$

これにより， E_2 はノイズの相関行列の対角成分だけに依存することがわかる．

3.4.3 周辺尤度最大化基準

教師付き学習とは異なり，実際には修復誤差最小化基準を画像修復に用いることはできない．なぜなら誤差を計算するために必要な原画像があれば，ノイズで汚れた劣化画像を修復する必要が無いからである．

そこで，ここでは原画像を直接用いなくてもハイパーパラメータの推定が行える周辺尤度最大化基準を議論する．式 (3.1),(3.8) の原画像とノイズの生成確率は既知である．それらを用いると，劣化画像の確率 $P(\tau)$ を計算することができる．

$$P(\tau) = \int d\mathbf{S} P_{\text{out}}(\tau|\mathbf{S}; \hat{R}) P(\mathbf{S}; \hat{\beta}, \hat{h}) \quad (3.55)$$

$P(\tau)$ は周辺尤度と呼ばれる．

劣化画像 τ が周辺尤度 $P(\tau)$ の最大値になるようにハイパーパラメータ $\hat{\beta}, \hat{h}, \hat{R}$ を決めることを，周辺尤度最大化と呼ぶ．

周辺尤度は式 (3.3),(3.9) を用いて次式で与えられる．

$$P(\tau) = \frac{Z_{\text{posterior}}(\hat{\beta}, \hat{h}, \hat{R})}{Z_{\text{noise}}(\hat{R}) Z_{\text{prior}}(\hat{\beta}, \hat{h})} \quad (3.56)$$

ただし，

$$\begin{aligned} & Z_{\text{posterior}}(\hat{\beta}, \hat{h}, \hat{R}) \\ &= \prod_{\mathbf{k}} \int d\tilde{S}_{\mathbf{k}} \exp(-\hat{H}_{\text{eff}}) \end{aligned} \quad (3.57)$$

$$= \pi^{\frac{N}{2}} \prod_{\mathbf{k}} \frac{1}{\sqrt{\hat{A}_{\mathbf{k}}}} \exp \left[-(\hat{B}_{\mathbf{k}} - \frac{\hat{B}_{\mathbf{k}}^2}{\hat{A}_{\mathbf{k}}}) |\tilde{\tau}_{\mathbf{k}}|^2 \right] \quad (3.58)$$

である．

\ln は単調増加関数であるので対数周辺尤度 $\ln P(\tau)$ を最大化するのは $P(\tau)$ を最大化するのと同じである．対数周辺尤度 $\ln P(\tau)$ は以下で与えられる．

$$\begin{aligned} \ln(P(\tau)) &= \ln \left(\frac{Z_{\text{posterior}}(\hat{\beta}, \hat{h}, \hat{R})}{Z_{\text{noise}}(\hat{R}) Z_{\text{prior}}(\hat{\beta}, \hat{h})} \right) \\ &= -\frac{1}{2} \sum_{\mathbf{k}} \ln \left(\hat{\beta} \tilde{G}_{\mathbf{k}} + \hat{h} + \frac{1}{2\hat{R}_{\mathbf{k}}} \right) \\ &\quad -\frac{N}{2} \ln(2\pi) - \frac{1}{2} \sum_{\mathbf{k}} \ln(\hat{R}_{\mathbf{k}}) \\ &\quad + \frac{1}{2} \sum_{\mathbf{k}} \ln(\hat{\beta} \tilde{G}_{\mathbf{k}} + \hat{h}) \end{aligned}$$

$$- \sum_{\mathbf{k}} \frac{(\hat{\beta}\tilde{G}_{\mathbf{k}} + \hat{h}) \frac{1}{2\hat{R}_{\mathbf{k}}}}{\hat{\beta}G_{\mathbf{k}} + \hat{h} + \frac{1}{2\hat{R}_{\mathbf{k}}}} \left| \tau_{\mathbf{k}} \right|^2 \quad (3.59)$$

3.5. シミュレーション

本節では，修復の際のハイパーパラメータを以下のように置くことにより，ノイズに空間的な相関が無い場合のモデルを用いて，空間的な相関があるノイズを修復することを考える．

$$\hat{R}_{i,j} = \hat{r}\delta_{i,j} \quad (3.60)$$

$$\hat{R}_{\mathbf{k}} = \hat{r} \quad (3.61)$$

この場合，式 (3.38) は，

$$\langle \sigma_j^{app} \rangle = \frac{1}{N} \sum_{\mathbf{k}} \sum_{\mathbf{i}} \frac{\tau_{\mathbf{i}} \frac{1}{2\hat{r}} \cos[\mathbf{k} \cdot (\mathbf{j} - \mathbf{i})]}{\hat{\beta}\tilde{G}_{\mathbf{k}} + \hat{h} + \frac{1}{2\hat{r}}} \quad (3.62)$$

となる．この式を式 (3.39) の $\langle \sigma_j \rangle$ に代入し， \hat{r} に関して，修復誤差を最小化した結果を図 3.1(B) に示す．

図 3.1 は $N = 16^2$ の二次元画像での解析結果と，同じサイズの人工画像 100 個を用いた数値計算の結果である．用いたパラメータは， $\beta = 1.0$ ， $h = 1.0$ ， $b = 0.75$ ， $\kappa = 3.0$ である．図の横軸は式 (3.11) の画像に重畳されるノイズの空間相関の大きさをあらわすパラメータ a であり， $a = 0$ が空間相関ない場合に相当する．

3.5.1 修復誤差最小化基準

(E) は修復前の誤差 E_2 をあらわしており，点線が解析解である式 (3.54) の結果を示し，実線が式 (3.53) を計算機実験で求めた場合の結果を示す．(E) の数値解が振動しているのは画素数が $N = 16^2$ ，画像サンプル数が 100 個と比較的少ないことによる．画素数を多くとる，もしくは画像サンプル数を増加させることで，(E) の数値解のゆらぎは小さくなり解析解に漸近していく．

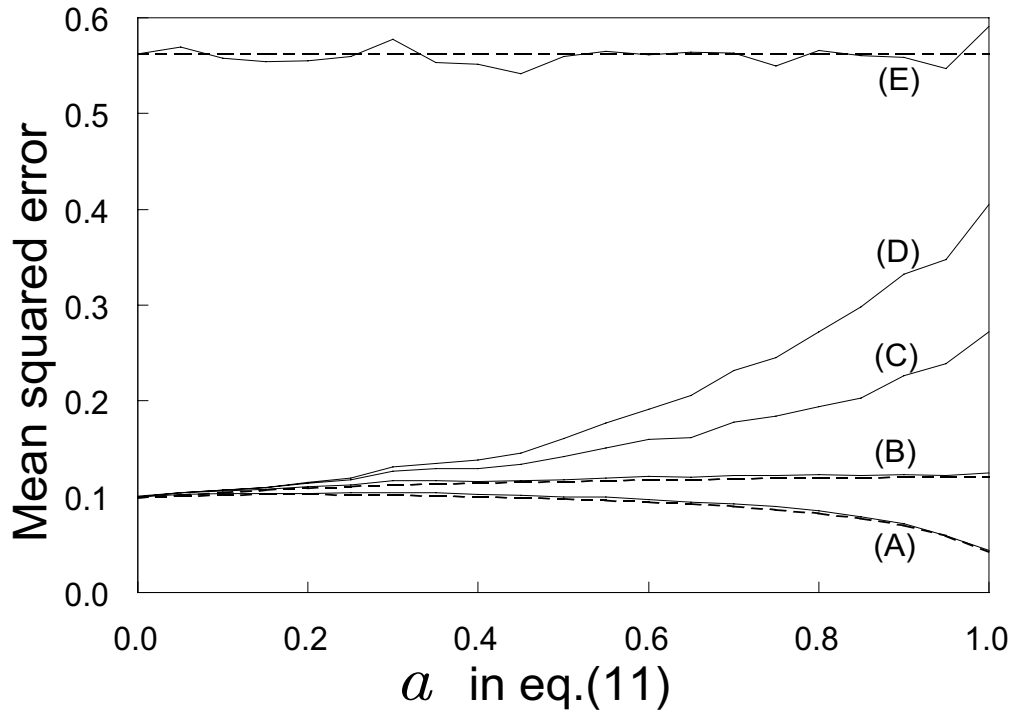


図 3.1 原画像との平均二乗誤差．横軸が式 (3.11) の a であり，縦軸が原画像との平均二乗誤差をあらわす．(A). 式 (3.52) で与えられる最適なデコードをした場合．(B). 修復誤差最小化により空間的な相関が無いモデルでデコードした場合．(C). 周辺尤度最大化により \hat{r} を推定した場合．(D). 周辺尤度最大化により $\hat{\beta}, \hat{h}, \hat{r}$ を推定した場合．(E). 式 (3.54) で与えられるデコード前の誤差．

(A) は式 (3.52) で与えられる，空間相関を持つノイズモデルに関して修復誤差が最小になるパラメータで修復した場合の修復誤差 $E_{1_{min}}$ をあらわす． $E_{1_{min}}$ は修復の限界値であるのでこれ以上の修復は理論的に不可能である．点線が解析解である式 (3.50) の結果を示し，実線が式 (3.39) を計算機実験で求めた場合の結果を示す．

(B) は空間相関 a が存在する場合に，式 (3.62) の空間相関が無いノイズモデルで修復を行なった際の修復誤差であり，式 (3.39) の \hat{r} に関して，修復誤差を最小化することにより得られた \hat{r} を用いて修復を行なった結果である．点線が解析解の結果であり，実線が計算機実験の結果である．修復誤差最小化によって得られた

\hat{r} を図 3.2 (B) に示す .

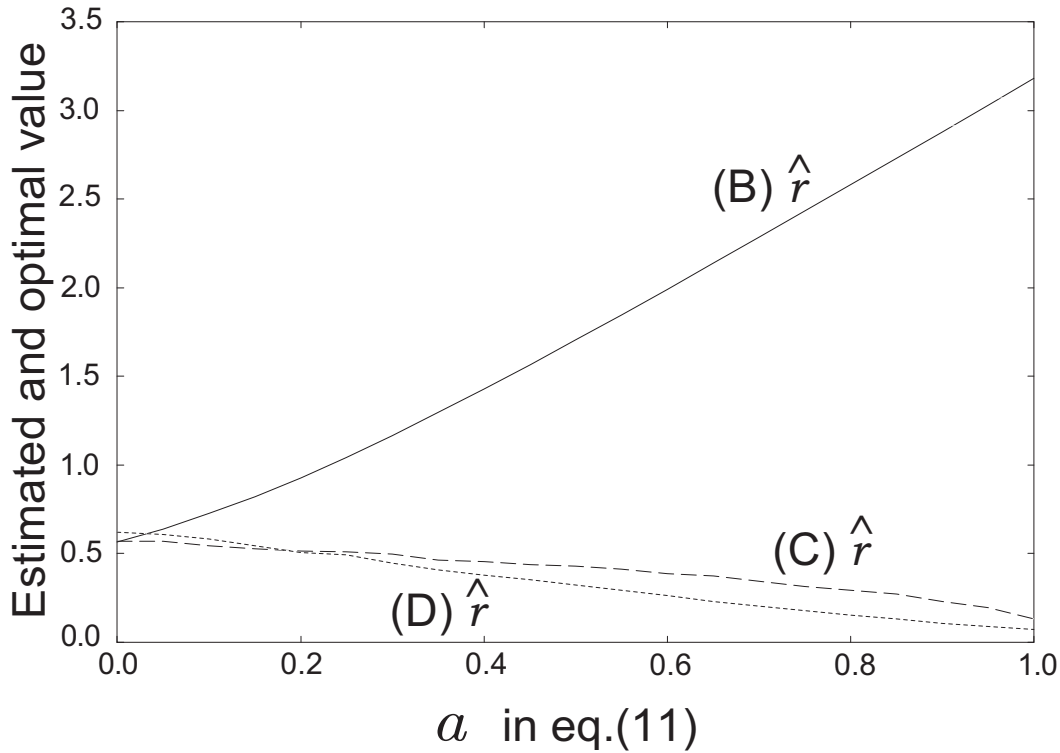


図 3.2 \hat{r} の推定結果をあらわすグラフ . 横軸は式 (3.11) の a であり , 縦軸は推定値 \hat{r} の平均値と修復誤差最小化により得られた最適値 \hat{r} をあらわす . 実線は図 3.1 の (B) に用いた \hat{r} , 波線は図 3.1 の (C) に用いた \hat{r} , 点線は図 3.1 の (D) に用いた \hat{r} である .

(E) の数値解が振動しているのに対して , (A),(B) の数値解はほとんど振動しておらず , 解析解と良く一致していることがわかる .

$a = 0$ の場合 , (A) と (B) は一致する . (B) はノイズに相関がないモデルによる修復誤差最小化基準での最適パフォーマンスをあらわすが , 前で述べたように , 実際は修復誤差最小化基準でハイパーパラメータを推定することはできない . そこで , 次に周辺尤度最大化基準でどの程度の修復誤差を小さくできるかを議論する .

3.5.2 周辺尤度最大化基準

式 (3.61) を式 (3.59) に代入すると ,

$$\begin{aligned}
 \ln(P_{app}(\tau)) &= -\frac{1}{2} \sum_{\mathbf{k}} \ln \left(\hat{\beta} \tilde{G}_{\mathbf{k}} + \hat{h} + \frac{1}{2\hat{r}} \right) \\
 &\quad - \frac{N}{2} \ln(2\pi) - \frac{N}{2} \ln(\hat{r}) \\
 &\quad + \frac{1}{2} \sum_{\mathbf{k}} \ln \left(\hat{\beta} \tilde{G}_{\mathbf{k}} + \hat{h} \right) \\
 &\quad - \sum_{\mathbf{k}} \frac{\left(\hat{\beta} \tilde{G}_{\mathbf{k}} + \hat{h} \right) \frac{1}{2\hat{r}}}{\hat{\beta} \tilde{G}_{\mathbf{k}} + \hat{h} + \frac{1}{2\hat{r}}} \left| \tilde{\tau}_{\mathbf{k}} \right|^2
 \end{aligned} \tag{3.63}$$

となる . $P_{app}(\tau)$ の極値条件は以下の非線形連立方程式に帰着される .

$$\begin{aligned}
 \frac{1}{\hat{\beta}} &= \left(\sum_{\mathbf{k}} \frac{\tilde{G}_{\mathbf{k}}}{\tilde{G}_{\mathbf{k}} + \frac{\hat{h}}{\hat{\beta}}} \right)^{-1} \\
 &\times \left[\sum_{\mathbf{k}} \frac{\tilde{G}_{\mathbf{k}}}{\hat{\beta} \tilde{G}_{\mathbf{k}} + \hat{h} + \frac{1}{2\hat{r}}} + \sum_{\mathbf{k}} \left| \tau_{\mathbf{k}} \right|^2 \frac{\tilde{G}_{\mathbf{k}}}{2\hat{r}^2 \left(\hat{\beta} \tilde{G}_{\mathbf{k}} + \hat{h} + \frac{1}{2\hat{r}} \right)^2} \right]
 \end{aligned} \tag{3.64}$$

$$\hat{r} = \frac{1}{2N} \sum_{\mathbf{k}} \frac{\tilde{G}_{\mathbf{k}}}{\hat{\beta} \tilde{G}_{\mathbf{k}} + \hat{h} + \frac{1}{2\hat{r}}} + \frac{1}{N} \sum_{\mathbf{k}} \left| \tilde{\tau}_{\mathbf{k}} \right|^2 \frac{\left(\hat{\beta} \tilde{G}_{\mathbf{k}} + \hat{h} \right)^2}{\left(\hat{\beta} \tilde{G}_{\mathbf{k}} + \hat{h} + \frac{1}{2\hat{r}} \right)^2} \tag{3.65}$$

$$\begin{aligned}
 \frac{1}{\hat{h}} &= \left(\sum_{\mathbf{k}} \frac{1}{1 + \frac{\hat{\beta} \tilde{G}_{\mathbf{k}}}{\hat{h}}} \right)^{-1} \\
 &\times \left[\sum_{\mathbf{k}} \frac{1}{\hat{\beta} \tilde{G}_{\mathbf{k}} + \hat{h} + \frac{1}{2\hat{r}}} + \sum_{\mathbf{k}} \left| \tau_{\mathbf{k}} \right|^2 \frac{1}{2\hat{r}^2 \left(\hat{\beta} \tilde{G}_{\mathbf{k}} + \hat{h} + \frac{1}{2\hat{r}} \right)^2} \right]
 \end{aligned} \tag{3.66}$$

実際にはこれらの式を繰り返し演算することで収束点が求まる .

得られた \hat{r} を用いて修復を行った結果を図 3.1(C) に示す . 画像の生成モデルは既知であるとし , 画像の生成確率のハイパーパラメータ β, h も既知であるとした . つまり修復の際のハイパーパラメータを $\hat{\beta} = \beta, \hat{h} = h$ とし , 相関のないノイズの分散 \hat{r} だけに関して周辺尤度を最大化することで , \hat{r} を求めた . 求めた \hat{r} を図 3.2(C) に示す . 図 3.2(C) では , ノイズの相関 a が大きくなるにつれ , 修復誤差最小化基準で得られた \hat{r} と周辺尤度最大化基準で得られた \hat{r} の差が大きくなっ

ている．その結果として，図 3.1(C) では，周辺尤度最大化で得られたハイパーパラメータを用いた際の修復誤差が大きくなっている．

これは，ノイズの生成モデルを正しく知らなければ，周辺尤度最大化によるハイパーパラメータ推定がうまくいかないことを意味している．つまり，今までの相関のないノイズモデルでは，空間的に相関のあるノイズモデルに対応できないことを示している．

これをさらに明らかにするために， \hat{r} だけではなく，画像の生成確率のハイパーパラメータ $\hat{\beta}, \hat{h}$ も同時に，周辺尤度最大化を用いて推定してみる．求めた \hat{r} を図 3.2(D) に， $\hat{\beta}, \hat{h}$ を図 3.3 に示す．これらのハイパーパラメータを用いて修復を行っ

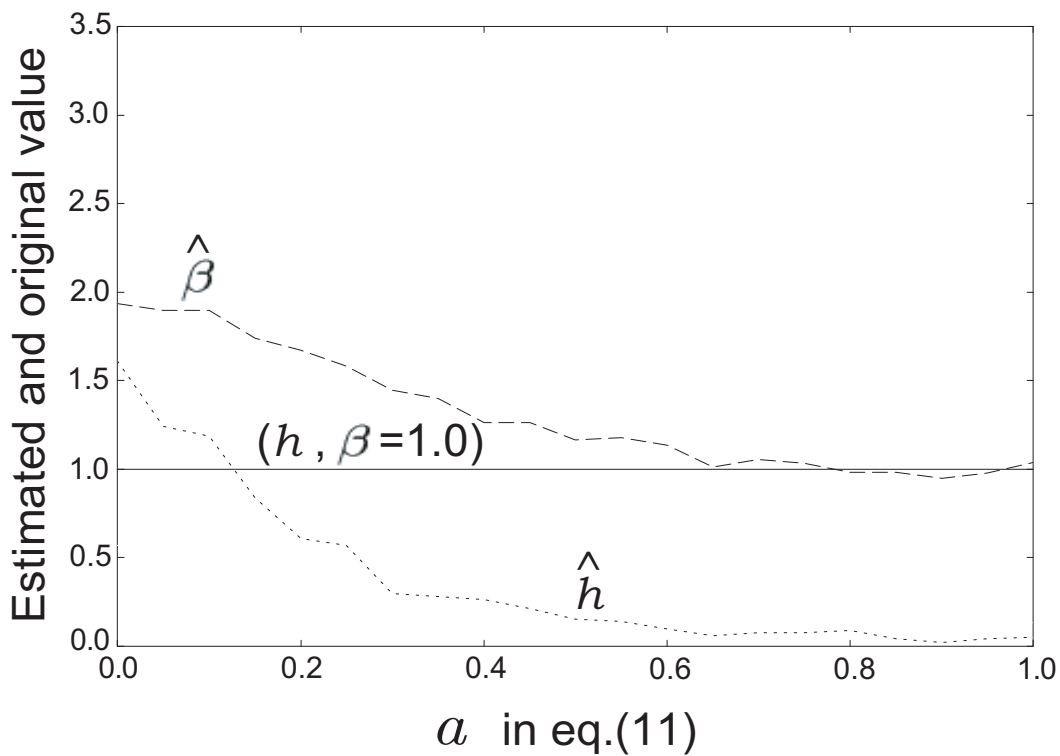


図 3.3 図 3.1 の (D) で用いた \hat{h} (点線)， $\hat{\beta}$ (波線) の平均値をあらわすグラフ．横軸は式 (3.11) の a であり，縦軸は推定値 $\hat{\beta}, \hat{h}$ の平均値と正しい値 β, h をあらわす．実線は正しい値 $h = \beta = 1.0$ である．

た結果を図 3.1(D) に示す．ノイズの分散 \hat{r} のみを推定した (C) の場合に比べて，

さらに修復誤差が大きくなっていることがわかる。

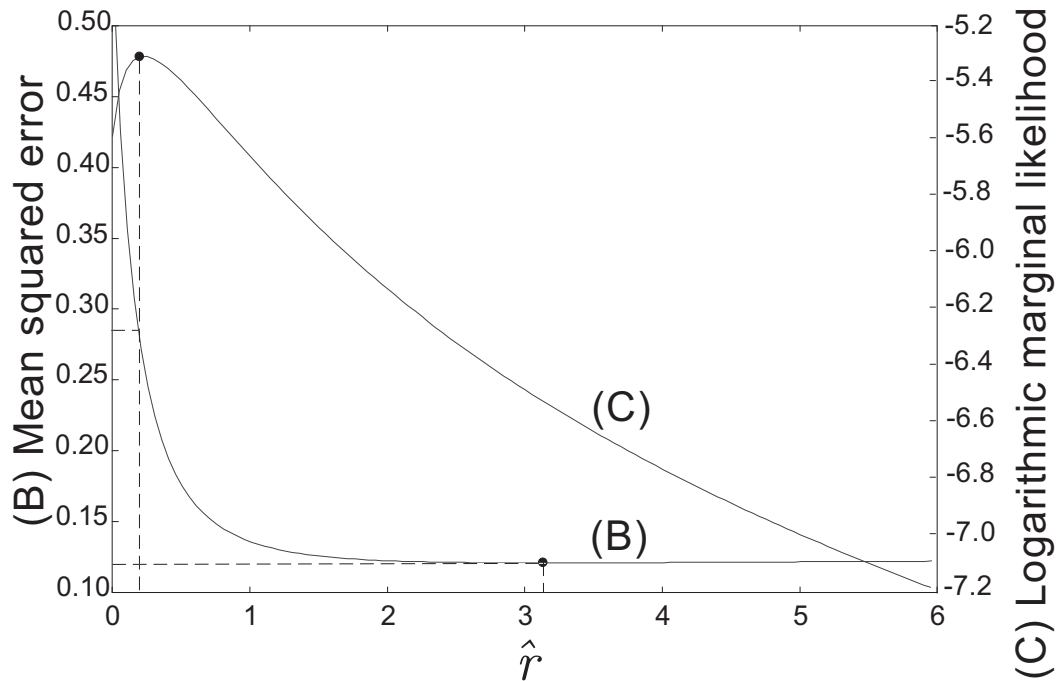


図 3.4 平均二乗誤差と対数周辺尤度のグラフ．横軸はハイパーパラメータ \hat{r} をあらわし，左の縦軸は式 (3.50) で与えられる平均二乗誤差 (B) であらわし，右の縦軸は式 (3.63) で与えられる対数周辺尤度 (C) をあらわす．用いたパラメータは $a = 1.0$, $\beta = 1.0$, $h = 1.0$, $\kappa = 3.0$ とした．(B) で用いた条件は図 3.1(B) の $a = 1.0$ における条件と同一であり，同様に，(C) で用いた条件は，は図 3.1(C) の $a = 1.0$ における条件と同一である．

次に，図 3.1(C) の対数周辺尤度が局所的な最大値を持たないことと，式 (3.66) により帰納的に求められた解がグローバルな最大値であることを示す．図 3.4(B) は式 (3.50) で与えられる平均二乗誤差 E_1 をプロットしたものであり，図 3.4(C) は式 (3.63) で与えられる対数周辺尤度 $\ln(P_{app}(\tau))$ をプロットしたものである．横軸は \hat{r} を示す．図 3.4(B) より， $\hat{r} = 3.2$ において E_1 が最小値 $E_1 = 0.12$ をとっていることが分かる．図 3.4(B) で用いられたパラメータは，図 3.1(B) の $a = 1.0$ の場合に用いたパラメータと同じである．図 3.4(C) は $\hat{r} = 0.2$ において対数周辺尤度が最大値 $E_1 = 0.28$ をとっていることを示している．図 3.4(C) で用いられてい

るパラメータは，図 3.1(C) の $a = 1.0$ の場合に用いたパラメータと同じである．重要なことは，対数周辺尤度である図 3.4(C) が極小値を持たないことである．

3.5.3 サンプル画像

次に，ガウス分布で生成した人工画像の修復例を見せる．まずハイパーパラメー

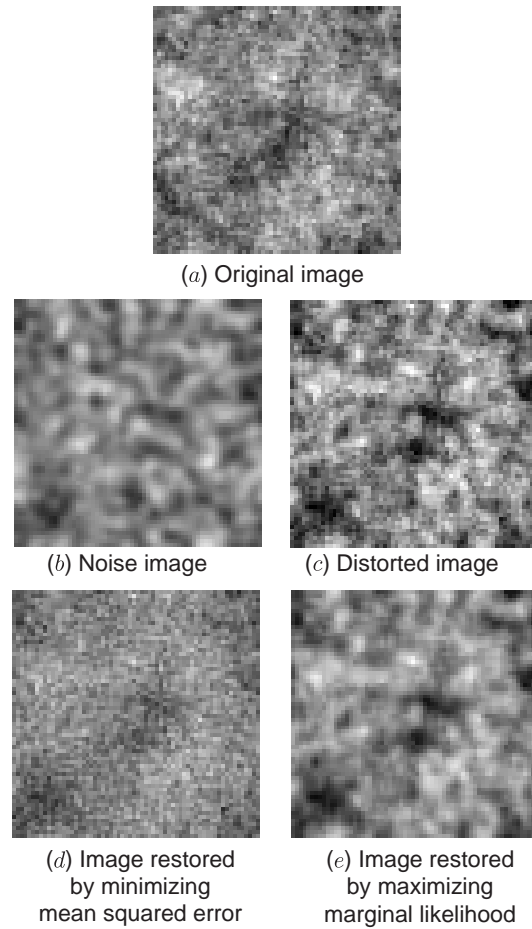


図 3.5 人工画像． $N = 64^2$ ．

タ $\beta = 0.5, h = 1.0^{-4}$ を用いて図 3.5(a) 原画像を生成した．それに $a = 1.0, b = 0.75, \kappa = 3.0$ のノイズ図 3.5(b) を重畳した画像が図 3.5(c) の劣化画像である．原画像と劣化画像の二乗誤差 E_2 は $E_2 = 0.57$ である．生成モデルと修復の際のモデ

ルが完全に一致しているとして、修復誤差最小化基準で全てのハイパーパラメータを最適化した場合の修復画像を図3.5(d)に示す。これは図3.1(A)に対応し、修復後の二乗誤差 E_1 は $E_1 = 0.27$ である。一方、図3.5(e)はノイズに相関が無いと仮定した場合に、周辺尤度最大化でハイパーパラメータを推定した場合の修復画像である。これは図3.1(D)に対応し、修復後の二乗誤差 E_1 は $E_1 = 0.45$ である。図3.5(d)は見た目にも原画像(a)に似ていいるが、図3.5(e)はむしろ劣化画像(c)に似ておりピントもずれたような画像になっている。視覚的にも(e)の修復誤差が大きくなっていることがわかる。

さらに、自然画像を用いた場合の結果を示す。図3.6は自然画像と、推定され

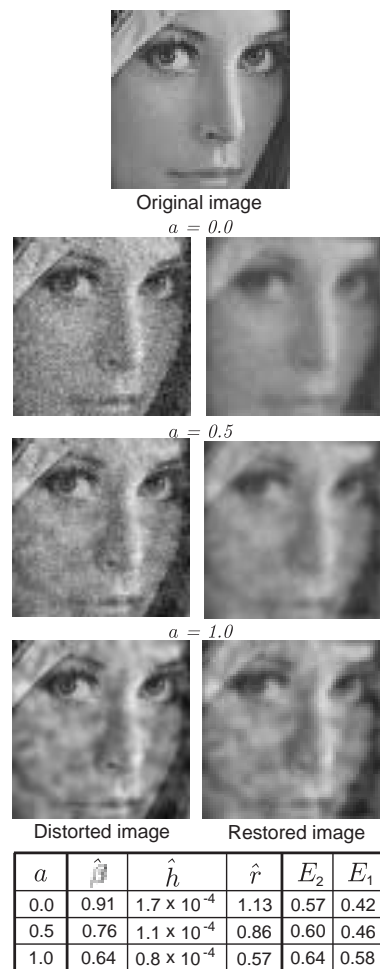


図 3.6 自然画像 . $N = 64^2$.

たパラメータの値である．ノイズパラメータとして $b = 0.75$ と $\kappa = 3.0$ を用い， a には $0.0, 0.5, 1.0$ の3つの値を用意した．左側の3つの画像は劣化画像であり，その右の画像は周辺尤度最大化によって $\hat{\beta}, \hat{h}, \hat{r}$ を推定することにより修復された画像である．このシミュレーション結果においても， a の増加に伴い E_1 が増加している．この傾向は，3つの修復画像においても反映されている． $a = 1.0$ における修復画像では，シミのように見えるノイズがまだ残っているのが確認できる．

3.5.4 ノイズモデルの特異点について

ノイズモデルの相関行列 (3.22) のハイパーパラメータは a, b, κ であるので，これらの推定値を $\hat{a}, \hat{b}, \hat{\kappa}$ と置くと推定された相関行列 $\hat{R}_{\mathbf{k}}$ は次のようになる．

$$\hat{R}_{\mathbf{k}} = (1 - \hat{a})\hat{b}^2 + \hat{a}\hat{b}^2 \sum_{\mathbf{l}} e^{-\frac{\mathbf{l}^2}{\hat{\kappa}^2}} \cos(\mathbf{k} \cdot \mathbf{l}) \quad (3.67)$$

ハイパーパラメータ $\hat{a}, \hat{b}, \hat{\kappa}$ を求めるには，

$$\begin{aligned} \ln(P(\boldsymbol{\tau})) &= -\frac{1}{2} \sum_{\mathbf{k}} \ln \left(\hat{\beta} \tilde{G}_{\mathbf{k}} + \hat{h} + \frac{1}{2\hat{R}_{\mathbf{k}}} \right) \\ &\quad - \frac{N}{2} \ln(2\pi) - \frac{1}{2} \sum_{\mathbf{k}} \ln \left(\hat{R}_{\mathbf{k}} \right) \\ &\quad + \frac{1}{2} \sum_{\mathbf{k}} \ln \left(\hat{\beta} \tilde{G}_{\mathbf{k}} + \hat{h} \right) \\ &\quad - \sum_{\mathbf{k}} \frac{\left(\hat{\beta} \tilde{G}_{\mathbf{k}} + \hat{h} \right) \frac{1}{2\hat{R}_{\mathbf{k}}}}{\hat{\beta} \tilde{G}_{\mathbf{k}} + \hat{h} + \frac{1}{2\hat{R}_{\mathbf{k}}}} \left| \boldsymbol{\tau}_{\mathbf{k}} \right|^2 \end{aligned} \quad (3.68)$$

を $\hat{a}, \hat{b}, \hat{\kappa}$ でそれぞれ偏微分することで漸化式を求め，その停留点を推定値とすればよい．漸化式は以下で与えられる．

$$\begin{aligned} \hat{a} &= \left[\sum_{\mathbf{k}} \left\{ \left(\frac{1}{\hat{a}} - 1 \right) \hat{b}^2 + \hat{b}^2 \sum_{\mathbf{l}} e^{-\frac{\mathbf{l}^2}{\hat{\kappa}^2}} \cos(\mathbf{k} \cdot \mathbf{l}) \right\} \frac{\partial}{\partial \hat{a}} Q_{\mathbf{k}} \right]^{-1} \\ &\quad \times \left[\frac{1}{2} \sum_{\mathbf{k}} \frac{\frac{\partial}{\partial \hat{a}} Q_{\mathbf{k}}}{\hat{\beta} \tilde{G}_{\mathbf{k}} + \hat{h} + \frac{1}{2\hat{R}_{\mathbf{k}}}} \right] \end{aligned}$$

$$+ \sum_{\mathbf{k}} \left(\hat{\beta} \tilde{G}_{\mathbf{k}} + \hat{h} \right) |\tau_{\mathbf{k}}|^2 \frac{\frac{\partial}{\partial \hat{a}} Q_{\mathbf{k}} \left(\hat{\beta} G_{\mathbf{k}} + \hat{h} \right)}{\left(\hat{\beta} G_{\mathbf{k}} + \hat{h} + \frac{1}{2\hat{R}_{\mathbf{k}}} \right)^2} \Big] \quad (3.69)$$

$$\begin{aligned} \hat{b}^2 &= \left[\sum_{\mathbf{k}} \left\{ (1 - \hat{a}) + \hat{a} \sum_{\mathbf{l}} e^{-\frac{l^2}{\hat{\kappa}^2}} \cos(\mathbf{k} \cdot \mathbf{l}) \right\} \frac{\partial}{\partial \hat{b}^2} Q_{\mathbf{k}} \right]^{-1} \\ &\times \left[\frac{1}{2} \sum_{\mathbf{k}} \frac{\frac{\partial}{\partial \hat{b}^2} Q_{\mathbf{k}}}{\hat{\beta} G_{\mathbf{k}} + \hat{h} + \frac{1}{2\hat{R}_{\mathbf{k}}}} \right. \\ &\quad \left. + \sum_{\mathbf{k}} \left(\hat{\beta} \tilde{G}_{\mathbf{k}} + \hat{h} \right) |\tau_{\mathbf{k}}|^2 \frac{\frac{\partial}{\partial \hat{b}^2} Q_{\mathbf{k}} \left(\hat{\beta} G_{\mathbf{k}} + \hat{h} \right)}{\left(\hat{\beta} G_{\mathbf{k}} + \hat{h} + \frac{1}{2\hat{R}_{\mathbf{k}}} \right)^2} \right] \quad (3.70) \end{aligned}$$

$$\begin{aligned} \hat{\kappa}^2 &= \left[-\frac{\hat{a}\hat{b}^2}{\hat{\kappa}^2} \sum_{\mathbf{k}} \frac{1}{R_{\mathbf{k}}} \sum_{\mathbf{l}} l^2 e^{-\frac{l^2}{\hat{\kappa}^2}} \cos(\mathbf{k} \cdot \mathbf{l}) \right] \\ &\times \left[\frac{1}{2} \sum_{\mathbf{k}} \frac{\frac{\partial}{\partial \hat{b}^2} Q_{\mathbf{k}}}{\hat{\beta} G_{\mathbf{k}} + \hat{h} + \frac{1}{2\hat{R}_{\mathbf{k}}}} \right. \\ &\quad \left. + \sum_{\mathbf{k}} \left(\hat{\beta} \tilde{G}_{\mathbf{k}} + \hat{h} \right) |\tau_{\mathbf{k}}|^2 \frac{\frac{\partial}{\partial \hat{\kappa}^2} Q_{\mathbf{k}} \left(\hat{\beta} G_{\mathbf{k}} + \hat{h} \right)}{\left(\hat{\beta} G_{\mathbf{k}} + \hat{h} + \frac{1}{2\hat{R}_{\mathbf{k}}} \right)^2} \right]^{-1} \quad (3.71) \end{aligned}$$

ただし,

$$\frac{\partial}{\partial \hat{a}} Q_{\mathbf{k}} = -\frac{1}{R_{\mathbf{k}}^2} \left(-\hat{b}^2 + \hat{b}^2 \sum_{\mathbf{l}} e^{-\frac{l^2}{\hat{\kappa}^2}} \cos(\mathbf{k} \cdot \mathbf{l}) \right) \quad (3.72)$$

$$\frac{\partial}{\partial \hat{b}^2} Q_{\mathbf{k}} = -\frac{1}{R_{\mathbf{k}}^2} \left(1 - \hat{a} + \hat{a} \sum_{\mathbf{l}} e^{-\frac{l^2}{\hat{\kappa}^2}} \cos(\mathbf{k} \cdot \mathbf{l}) \right) \quad (3.73)$$

$$\frac{\partial}{\partial \hat{\kappa}^2} Q_{\mathbf{k}} = -\frac{1}{R_{\mathbf{k}}^2} \left(\hat{a}\hat{b}^2 \frac{1}{\hat{\kappa}^4} \sum_{\mathbf{l}} e^{-\frac{l^2}{\hat{\kappa}^2}} \cos(\mathbf{k} \cdot \mathbf{l}) \right) \quad (3.74)$$

である．

ところが，式 (3.67) には特異点があり，漸化式 (3.69)，(3.70)，(3.71) から停留点を求めるのは困難である．例えば，元の a の値が $a \ll 1$ の場合，式 (3.67) の第二項の影響が極めて少なくなる．この場合 $\hat{\kappa}$ がどのような値をとっても良いことになり，漸化式 (3.69)，(3.70)，(3.71) を数値計算をして行く途中で値が発散してしまう．つまり $\hat{a} = 0.0$ は漸化式 (3.71) にとって特異点となっている．

また， $\hat{a} \ll 1$ としない場合，例えば元の値を $a = 0.5$ とした場合においても $\hat{a} = a$ ， $\hat{b} = b$ ， $\hat{\kappa}^2 = \kappa^2$ となる点は漸化式の安定点とはならないことが多い．その様子を図 3.7 に示す．

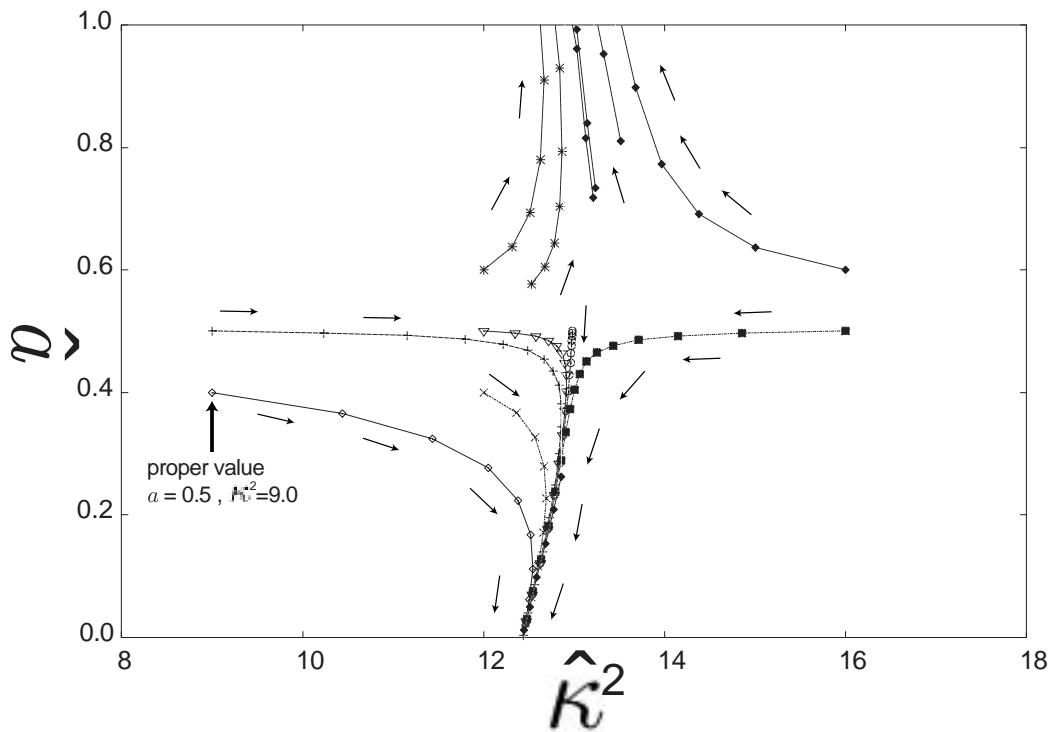


図 3.7 元の値を $a = 0.5$ ， $b = 0.75$ ， $\kappa^2 = 9.0$ とした場合に，漸化式 (3.69)，(3.70) において \hat{a} ， $\hat{\kappa}^2$ を繰り返し演算した場合の相図．横軸に $\hat{\kappa}^2$ ，縦軸に \hat{a} としている． $\hat{b} = 0.75$ と固定した．正しい値 $\hat{a} = 0.5$ ， $\hat{\kappa}^2 = 9.0$ を代入した場合でも， \hat{a} がその定義域を超えて発散してしまう様子が見られる．

図 3.7 では元の値を $a = 0.5$ ， $b = 0.75$ ， $\kappa^2 = 9.0$ とし， $\hat{b} = 0.75$ を固定するこ

とで，漸化式 (3.69)，(3.70) を用いて \hat{a} ， $\hat{\kappa}^2$ を繰り返し演算した場合の相図である． a の定義域は $1 \leq a \leq 0$ であるにもかかわらず， $a < 0$ となったり， $a > 1$ となっていく様子が見られる．正しい値に収束することは決して無く，いずれ数値計算はオーバーフローを起こす．

ただし， $\hat{a} = a$ ， $\hat{b} = b$ として固定し， $\hat{\kappa}$ だけを求める場合は， $\hat{\kappa} = \kappa$ が漸化式 (3.71) の収束点となり，旨く行くことが分かっている．

3.6. 議論

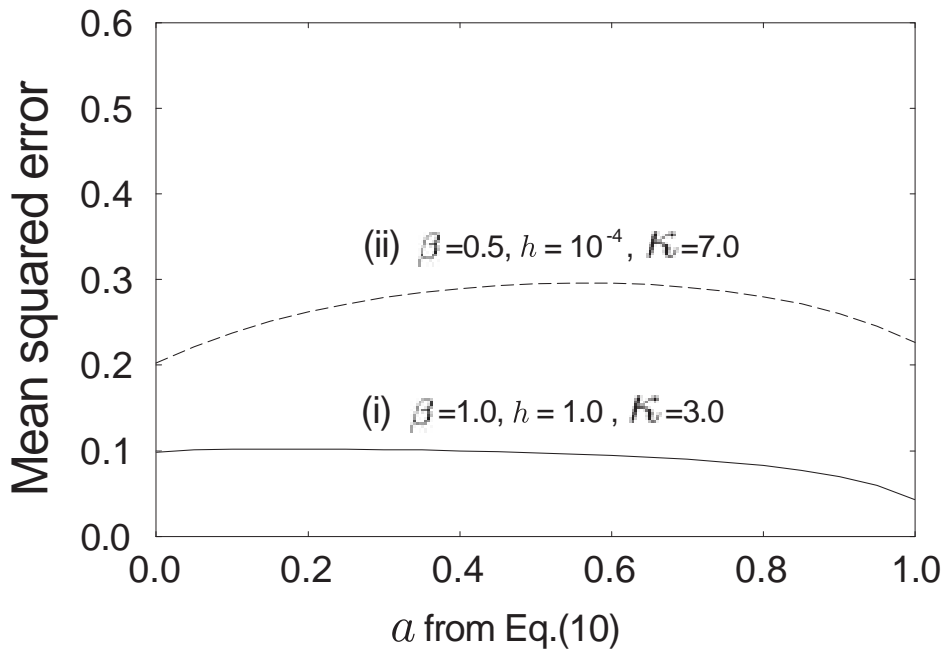


図 3.8 原画像と修復画像の平均二乗誤差．横軸は式 (3.11) の a を示し，縦軸は平均二乗誤差をあらわす．式 (3.52) で与えられる修復誤差最小化基準を用いてデコードしている．パラメータ $b = 0.75$ と $N = 64^2$ は固定している．実線 (i) では， $\beta = 1.0, h = 1.0, \kappa = 3.0$ を用い，点線 (ii) では， $\beta = 0.5, h = 10^{-4}, \kappa = 7.0$ を用いている．

ここでは，なぜ図 3.1(A) が $a = 1.0$ に近づくにつれて減少していくかについて

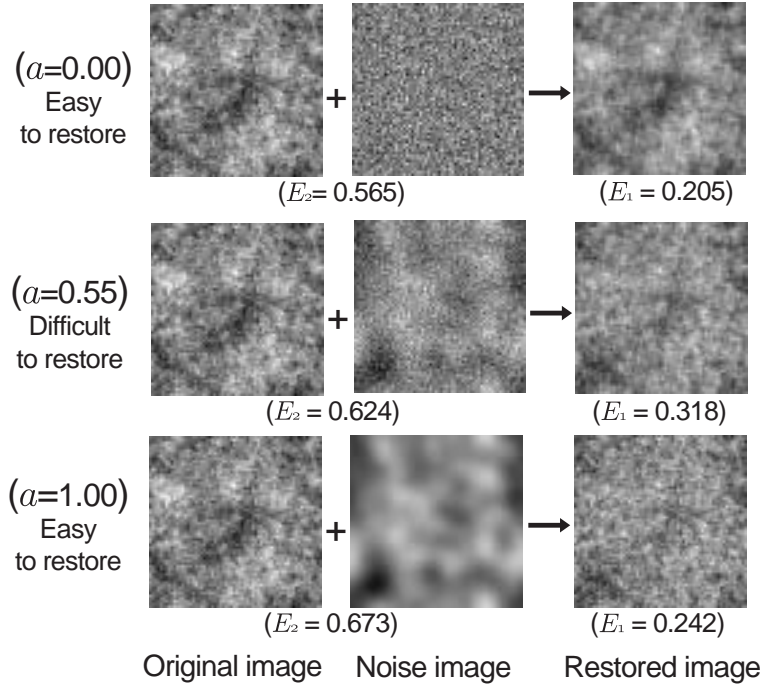


図 3.9 図 3.8 (ii) の典型例 . $N = 64^2$ とした . 原画像で用いたパラメータは $\beta = 0.5$ と $h = 10^{-4}$ であり , 3つの原画像は同一である . ノイズ画像で用いられたパラメータは $\beta = 0.5, h = 10^{-4}, \kappa = 7.0, b = 0.75$ であり , a のみが変わっている . $a = 0.55$ のとき , 原画像とノイズ画像が見た目に似ており , 修復が最もうまく行っていないのが分かる .

議論する . 実はこの現象はいかなるパラメータセットにおいても普遍的に生じる .

その一例を図 3.8(ii) に示す . 図 3.8(ii) では図 3.1(A) とは異なるパラメータセット $\beta = 0.5, h = 10^{-4}, \kappa = 7.0$ が用いられている . 参考のために図 3.8 には図 3.1(A) と同じパラメータセットを用いた場合のグラフを図 3.8(i) に掲載する . ただし , N を $N = 64^2$ に変更している . しかしながら , 図 3.1(A) と図 3.8(i) は同一になっている . つまり図 3.1(A) は N に依存しない . これは式 (3.7) の周期境界条件を課しているためである . 原画像の生成モデルには , 式 (3.1) により与えられる最近接相互作用を持つガウス分布を用いる . ノイズの生成モデルには , 式 (3.8) により与えられる相互作用がガウス関数的に減少するガウス分布を用いる . $a = 1.0$ では , 原画像とノイズの生成モデルの違いが大きくなり , この違いが画像修復をう

まく生かせていると考えられる。

そこでこのことについて検証してみる。図 3.9 に図 3.8(ii) の典型例を示す。図 3.9($a = 0.55$) では最も修復するのが難しくなっており、これは図 3.8(ii) の結果と一致する。この場合、原画像とノイズ画像が見た目に似ている。もっと具体的に言えば、2つの画像の相関長が似ている。したがって、劣化画像の修復を成功させるためには、原画像とノイズ画像の統計的性質の違いが必要になる。

ここで、図 3.1(A) がなぜ $a = 1.0$ に近づくにつれ減少するかについてのまとめを行う。原画像とノイズ画像の統計的性質、特に相関長の違いが大きければ大きいほど劣化画像の修復が容易になり、 a が 1.0 に近づくにつれ修復誤差は減少する。逆に、原画像とノイズ画像の相関長の違いが小さい場合は、2つの画像を区別することが難しくなり、劣化画像を修復するのが難しくなる。この修復が困難な場合の典型例が図 3.8(ii) の $a = 0.55$ のときである。劣化画像の修復を成功させるためには、原画像とノイズ画像の相関長の長さが異ならなくてはならない。図 3.9 の $a = 0.00$ と $a = 1.00$ の場合は、これら 2つの画像の相関長が異なるので修復が旨く行っている。

3.7. 本章のむすび

本章では空間的な相関ノイズを持つ場合の画像修復をベイズ推定の立場から議論した。フーリエ変換を用いることで相関行列を対角化し、修復の限界値等の統計量を解析的に評価した。そこで、修復画像の期待値を求め、修復誤差最小化基準によりガウス分布を決めるハイパーパラメータを求めた。

次に、修復の際に空間的な相関を持たないノイズを仮定した場合に、どのような修復結果を得られるかを議論した。ハイパーパラメータ推定の基準として、修復誤差最小化基準と周辺尤度最大化基準を用いた。ノイズの空間相関が大きくなると、修復誤差最小化基準と周辺尤度最大化基準で得られたハイパーパラメータの値の差が大きくなる。その結果として、周辺尤度最大化で得られたハイパーパラメータを用いた際の修復誤差が大きくなっていることがわかった。その理由は、周辺尤度最大化にはシグナルである原画像とノイズを分離するメカニズムが含ま

れてないため、ノイズの分散が推定できなくなり修復誤差が大きくなったためであると考えられる。

最後に、修復がうまく行く場合と行かない場合の理由について議論した。その結果、原画像とノイズ画像の統計的性質（特に相関長）が異なっている場合は修復がうまく行き、似ている場合はうまく行かないことが分かった。

3.8. 本章の付録

3.8.1 κ について

図 3.8 では, a を調節することでノイズ画像のシミの大きさを変化させているが, κ を変化させることでもノイズ画像のシミの大きさを変化させることはできる. 図 3.10 に a, b を $a = 1.0, b = 0.75$ とそれぞれ固定し, κ を $\kappa = 0.5, 3.0, 7.0$ と変化させた場合におけるノイズ画像を見せる. 図 3.10 より, 図 3.8 のノイズ画像

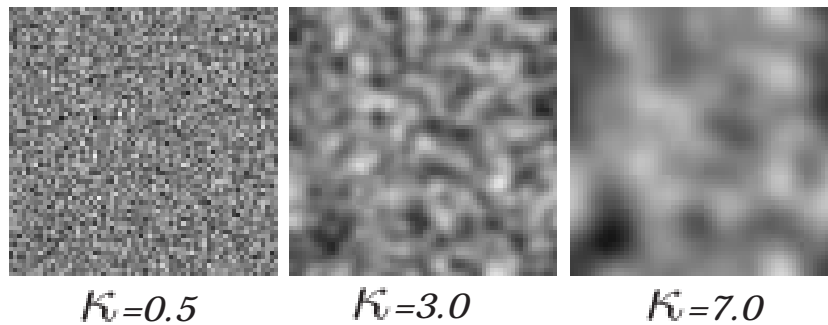


図 3.10 様々な κ に対するノイズ画像. $a = 0.0, b = 0.75$ と固定し, κ を $\kappa = 0.5, 3.0, 7.0$ と変化させた場合.

と同様にシミの大きさが大きくなっているのが確認できる. ちなみに図 3.10 の $\kappa = 7.0$ のノイズ画像は図 3.8 の $a = 1.0$ におけるノイズ画像と同じである. このように, a と同様に κ を小さくすることでシミを細かくすることができ, 逆に, 大きくすることでシミを荒くすることができるが, 本章では a を変化させることに主点をおいて議論を行った.

$a = 1.0$ と固定することで式 (3.11) で与えられる相関行列の右辺第一項を無視することができ, パラメータを二つに抑えることができる. よって, 実際の工学的な応用を考える場合は $a = 1.0$ と固定し, κ と b のみでノイズを議論するのが好ましい.

3.8.2 式 (3.20) の導出法

ここでは，式 (3.20),(3.21) を導出するにおいて必要になる，式 (3.2) のフーリエ変換による対角化について解説する．式 (3.2) の右辺を展開すると，

$$\xi^T(\beta G + I)\xi = \beta \xi^T G \xi + \xi^T \xi \quad (3.75)$$

となるので，まずは式 (3.75) の右辺第二項 $\xi^T \xi$ をフーリエ表記する．

$$\xi^T \xi \quad (3.76)$$

$$= \sum_j \xi_j^2 \quad (3.77)$$

$$= \sum_j \left(\frac{1}{\sqrt{N}} \sum_{k_1} \tilde{\xi}_{k_1} e^{i k_1 \cdot j} \right) \times \left(\frac{1}{\sqrt{N}} \sum_{k_2} \tilde{\xi}_{k_2} e^{i k_2 \cdot j} \right) \quad (3.78)$$

$$= \frac{1}{N} \sum_j \sum_{k_1} \sum_{k_2} \tilde{\xi}_{k_1} \tilde{\xi}_{k_2} e^{i(k_1+k_2) \cdot j} \quad (3.79)$$

$$= \frac{1}{N} \sum_{k_1} \sum_{k_2} \delta_{(k_1+k_2),0} \tilde{\xi}_{k_1} \tilde{\xi}_{k_2} \sum_j \quad (3.80)$$

$$= \sum_{k_1} \sum_{k_2} \delta_{k_1, -k_2} \tilde{\xi}_{k_1} \tilde{\xi}_{k_2} \quad (3.81)$$

$$= \sum_k \tilde{\xi}_k \tilde{\xi}_{-k} \quad (3.82)$$

次に，式 (3.75) の第一項 $\beta \xi^T G \xi$ をフーリエ表記する．

$\xi^T G \xi$ に式 (3.4) を代入すると，

$$\xi^T G \xi = \sum_j \sum_{\delta} (\xi_j - \xi_{j+\delta})^2 \quad (3.83)$$

となる．よって，これを逆フーリエ変換する．

$$\sum_j \sum_{\delta} (\xi_j - \xi_{j+\delta})^2 \quad (3.84)$$

$$\begin{aligned}
&= \sum_{\mathbf{j}} \sum_{\boldsymbol{\delta}} \\
&\quad \times \left[\sum_{\mathbf{k}_1} \frac{1}{\sqrt{N}} (\tilde{\xi}_{\mathbf{k}_1} e^{i\mathbf{k}_1 \cdot \mathbf{j}} - \tilde{\xi}_{\mathbf{k}_1} e^{i\mathbf{k}_1 \cdot (\mathbf{j} + \boldsymbol{\delta})}) \right] \\
&\quad \times \left[\sum_{\mathbf{k}_2} \frac{1}{\sqrt{N}} (\tilde{\xi}_{\mathbf{k}_2} e^{i\mathbf{k}_2 \cdot \mathbf{j}} - \tilde{\xi}_{\mathbf{k}_2} e^{i\mathbf{k}_2 \cdot (\mathbf{j} + \boldsymbol{\delta})}) \right]
\end{aligned} \tag{3.85}$$

$$\begin{aligned}
&= \frac{1}{N} \sum_{\mathbf{j}} \sum_{\boldsymbol{\delta}} \sum_{\mathbf{k}_1} \sum_{\mathbf{k}_2} \tilde{\xi}_{\mathbf{k}_1} \tilde{\xi}_{\mathbf{k}_2} \\
&\quad \times \left[e^{i(\mathbf{k}_1 + \mathbf{k}_2) \cdot \mathbf{j}} + e^{i(\mathbf{k}_1 + \mathbf{k}_2) \cdot (\mathbf{j} + \boldsymbol{\delta})} \right. \\
&\quad \left. - e^{i\mathbf{k}_1 \cdot (\mathbf{j} + \boldsymbol{\delta})} e^{i\mathbf{k}_2 \cdot \mathbf{j}} - e^{i\mathbf{k}_2 \cdot (\mathbf{j} + \boldsymbol{\delta})} e^{i\mathbf{k}_1 \cdot \mathbf{j}} \right]
\end{aligned} \tag{3.86}$$

$$\begin{aligned}
&= \frac{1}{N} \sum_{\mathbf{k}} \sum_{\boldsymbol{\delta}} \tilde{\xi}_{\mathbf{k}} \tilde{\xi}_{-\mathbf{k}} \\
&\quad \times \left[N + N - N e^{i\mathbf{k} \cdot \boldsymbol{\delta}} - N e^{-i\mathbf{k} \cdot \boldsymbol{\delta}} \right]
\end{aligned} \tag{3.87}$$

$$= \sum_{\mathbf{k}} \tilde{\xi}_{\mathbf{k}} \tilde{\xi}_{-\mathbf{k}} \sum_{\boldsymbol{\delta}} [2 - 2 \cos(\mathbf{k} \cdot \boldsymbol{\delta})] \tag{3.88}$$

$$= \sum_{\mathbf{k}} \tilde{\xi}_{\mathbf{k}} \tilde{G}_{\mathbf{k}} \tilde{\xi}_{-\mathbf{k}} \tag{3.89}$$

ただし，

$$\tilde{G}_{\mathbf{k}} = \sum_{\boldsymbol{\delta}} [2 - 2 \cos(\mathbf{k} \cdot \boldsymbol{\delta})] \tag{3.90}$$

とおいた．式 (3.83) と式 (3.89) を比較すれば明らかなように， $G_{\mathbf{k}}$ は $G_{i,j}$ をフーリエ表記したものであり， $G_{\mathbf{k}}$ は格子グリーン関数と呼ばれる [16]．

よって，式 (3.2) は次のように求まる．

$$H(\boldsymbol{\xi}, \beta, h) = \sum_{\mathbf{k}} \tilde{\xi}_{\mathbf{k}} \tilde{G}_{\mathbf{k}} \tilde{\xi}_{-\mathbf{k}} + \sum_{\mathbf{k}} \tilde{\xi}_{\mathbf{k}} \tilde{\xi}_{-\mathbf{k}} \tag{3.91}$$

$$= \sum_{\mathbf{k}} (\beta \tilde{G}_{\mathbf{k}} + h) \tilde{\xi}_{\mathbf{k}} \tilde{\xi}_{-\mathbf{k}} \tag{3.92}$$

これを式 (3.1) に代入することで, 式 (3.20) が求まる .

第4章 ハイブリッド型連想記憶モデル

4.1. この章のはじめに

2.3 節において、連想記憶モデルには偽記憶へ収束する問題があることについて述べたが、本章では Hopfield 型の連想記憶モデルを拡張することでハイブリッド型の学習を可能にした、偽記憶に収束しにくいモデルの提案を行う。

ある記憶事象の一部だけが与えられており、それをキーとして全体を想起するような心的過程があることは否定できない。このような心的過程を神経回路モデルを用いてモデル化したものを、連想記憶モデルと呼ぶ。実際に脳の海馬 CA3 ニューロン間の再帰的結合には、連想記憶に近い機能があることがわかりつつある [41]。ただし本章で提案するモデルは、現時点で海馬のネットワークに対応付けようとするものではない。

Hopfield のオリジナルの CAM(Content Addressable Memory)[37] は、全ニューロンによる相互結合のネットワークである。このモデルにはさらに二つの特徴があった。

- (1) 結合荷重行列が対称結合である。
- (2) 全ニューロンの状態が想起するパターンの表現に対応している。

上記 (1) は想起をエネルギー緩和過程ととらえるために都合が良いが、(1)(2) 共にネットワークに強い制限を加えていることになる。海馬のモデルとしても必然性はない。本章では上記 (1)(2) の仮定を除いたモデルを提案する。パターン表記部 (可視素子) 以外に、連想処理専用の隠れ素子を付け加えたモデルを考えるこ

とによって (2) の条件を除く．この隠れ素子の導入により，モデルの能力の向上が期待できる．

可視空間と隠れ空間（後述）の関係をどのように決めるかという問題は重要であるが，本章では可視空間から隠れ空間へ何らかの写像が行われていると仮定する．これにより，この写像を決める必要がある．ここで，多対1写像と1対1写像の二つの写像を挙げることができるが，後者の1対1写像を用いることにした．本章では最も簡単な1対1写像として線形写像を考える．

線形写像はKimuraら [30] が提案したネットワークでも用いられており，彼らはそのネットワークを使って軌道近似を行っている．本章では，Kimuraらのネットワークを拡張したネットワークにより，自己連想記憶モデルへの適用を行う．

可視部と隠れ部の写像に線形写像を用いることにより，可視部 隠れ部 可視部の一連の再帰的結合を，2層の階層型パーセプトロン（MLP）と見なすことができ，可視部と隠れ部の結合の学習に誤差逆伝播法を用いることができる．結果として可視部と隠れ部の結合は非対称なものとなり，上記 (1) の条件は満たされない．一方で，可視部の相互結合にはヘップ則を用いる．

一般に，誤差逆伝播法は教師付き学習であるが，上記の可視部と隠れ部の結合の学習には，記憶パターンのみを用いるので，その誤差逆伝播法は教師なし学習と見なせる．可視部の学習則であるヘップ則も教師なし学習であることから，使用した学習法はハイブリッド型の教師なし学習である．

ホップフィールドモデルの場合，一般に素子の数に対して偽記憶が指数関数的に増加することが知られている [35]．本章ではシミュレーションを用いることにより，提案する自己連想記憶モデルが，隠れ素子を持たない相互結合型ネットワーク [26],[33]-[36],[37],[29],[27] や誤差逆伝播法 [32],[38] を用いた階層型ネットワークによる自己連想記憶モデルに比べ，大きな引き込み領域を持ち，偽記憶を出力しにくいモデルであることを示す．

4.2. ネットワークの定式化

本節ではまず，隠れ素子付きネットワークを一つの相互結合型ネットワークと見なすことにより定式化を行う．時刻 t におけるネットワークの状態変数をベクトル $\mathbf{u}(t)$ によりあらわす．素子の総数を M とすれば，このベクトルの次元は M に等しく，その要素 $u_j(t)$, ($1 \leq j \leq M$) は各素子の内部状態をあらわしている．この状態変数 $\mathbf{u}(t)$ を更新する状態方程式は次の微分方程式によりあらわされる．

$$\tau \frac{d\mathbf{u}}{dt} = -\mathbf{u} + W\mathbf{g}(\mathbf{u}) + \mathbf{b} \quad (4.1)$$

ここで， $M \times M$ 次行列 W の成分である w_{ij} は j 番目から i 番目の素子への結合荷重値をあらわす． $g_i(\cdot)$ は i 番目の素子の出力関数， τ は時定数， \mathbf{b} は M 次元バイアスである．状態変数 $\mathbf{u}(t)$ が動く M 次元の空間を状態空間と呼ぶ．

4.3. 隠れ素子付き相互結合型ネットワーク

相互結合型ネットワークを隠れ部と可視部に分割し，隠れ部に属する素子を隠れ素子，可視部に属する素子を可視素子と呼ぶことにする．隠れ部の概念を図 4.1 にあらわす．アルファベットの「A」が表示されている部分が可視部，ランダムなパターンが表示されている部分が隠れ部である．可視部および隠れ部に属しているすべての素子が，その所属に関係なく互いに結合している．可視部を出力部として用い，隠れ部を処理専用のネットワークとして用いる．ボルツマンマシン [36],[32] や，その平均場近似 [31] と同様に，処理専用の隠れ部を導入することにより処理能力の向上が期待できる．

可視素子の数を N ，隠れ素子の数を R とするとき ($M = N + R$)，全素子の状態変数

$$\mathbf{u}(t) = (u_1(t), \dots, u_N(t), u_{N+1}(t), \dots, u_M(t))^T \quad (4.2)$$

は次に示すように，可視変数 $\mathbf{x}(t)$ と隠れ変数 $\mathbf{y}(t)$ に分けて表記される．

$$\begin{aligned} \mathbf{x}(t) &= (x_1(t), \dots, x_N(t))^T = (u_1(t), \dots, u_N(t))^T \\ \mathbf{y}(t) &= (y_1(t), \dots, y_R(t))^T = (u_{N+1}(t), \dots, u_M(t))^T \end{aligned} \quad (4.3)$$

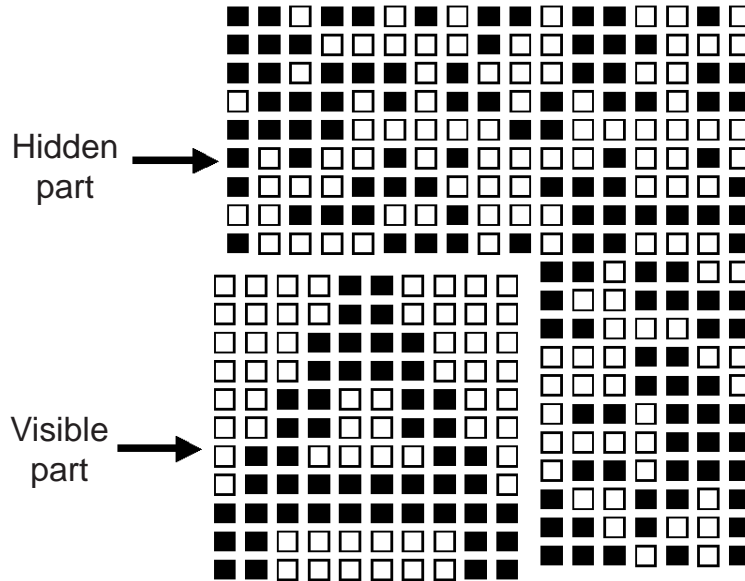


図 4.1 隠れ部の導入の様子をあらわす図 .

本章では，可視変数 $x(t)$ が動く N 次元空間を可視空間，隠れ変数 $y(t)$ が動く R 次元空間を隠れ空間と呼ぶ .

図 4.2 に $N = 2, R = 1$ の場合における状態変数軌道の様子を示す . 可視部および隠れ部の状態は，状態空間からそれぞれ可視空間および隠れ空間に対する射影により求まる .

式 (4.1) における結合荷重行列 $W (M \times M \text{ 次})$ を以下のように分割して表記する .

$$W = \begin{pmatrix} W^{XX} & W^{XY} \\ W^{YX} & W^{YY} \end{pmatrix} \quad (4.4)$$

$W^{XX} (N \times N \text{ 次})$ は可視部内の結合荷重行列， $W^{XY} (N \times R \text{ 次})$ は隠れ部から可視部への結合荷重行列， $W^{YX} (R \times N \text{ 次})$ は可視部から隠れ部への結合荷重行列， $W^{YY} (R \times R \text{ 次})$ は隠れ部内の結合荷重行列である . 式 (4.3), (4.4) を用いると，式 (4.1) を以下に示すように可視部と隠れ部に分割してあらわせる .

$$\tau \frac{d\mathbf{x}}{dt} = -\mathbf{x} + W^{XX} \mathbf{g}^X(\mathbf{x}) + W^{XY} \mathbf{g}^Y(\mathbf{y}) + \mathbf{b}^X \quad (4.5)$$

$$\tau \frac{d\mathbf{y}}{dt} = -\mathbf{y} + W^{YX} \mathbf{g}^X(\mathbf{x}) + W^{YY} \mathbf{g}^Y(\mathbf{y}) + \mathbf{b}^Y \quad (4.6)$$

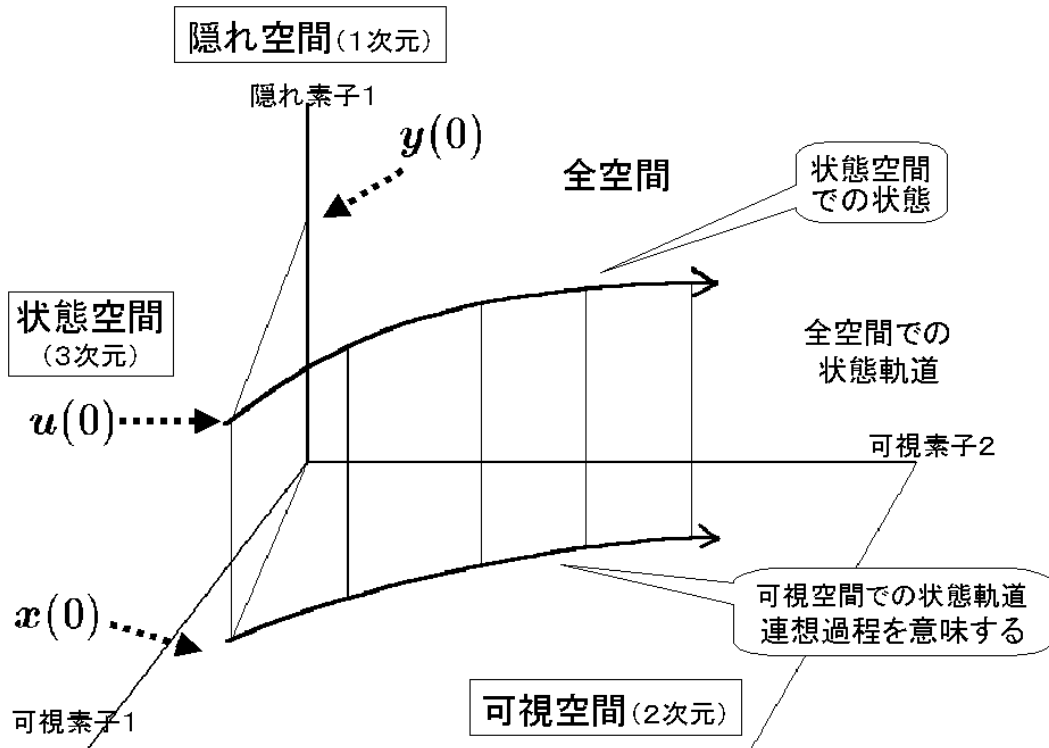


図 4.2 状態変数の軌道．可視素子の数 $N = 2$ ，隠れ素子の数 $R = 1$ ．可視素子 1 と可視素子 2 の可視変数が動く部分空間が可視空間 (2 次元) になり，隠れ素子 1 の隠れ変数が動く部分空間が隠れ空間 (1 次元) になる．

ここで， b^X ， b^Y はそれぞれ，可視部のバイアスベクトル，隠れ部のバイアスベクトルをあらわしている．また， $g^X(\cdot)$ ， $g^Y(\cdot)$ はそれぞれ，可視部の出力関数ベクトル，隠れ部の出力関数ベクトルをあらわしている．

$$\begin{aligned} \mathbf{g}^X(\mathbf{x}) &= [g_1^X(x_1(t)), \dots, g_N^X(x_N(t))]^T, \\ \mathbf{g}^Y(\mathbf{y}) &= [g_1^Y(y_1(t)), \dots, g_R^Y(y_R(t))]^T \end{aligned} \quad (4.7)$$

4.3.1 モデルの拡張

式 (4.5),(4.6) では素子の出力である $\mathbf{g}^X(\mathbf{x}), \mathbf{g}^Y(\mathbf{y})$ から内部状態変数への作用を線形であるとしている．ここで，非線形的作用を仮定し，式 (4.5),(4.6) 中の $W^{XX}\mathbf{g}^X(\mathbf{x})$ ， $W^{XY}\mathbf{g}^Y(\mathbf{y})$ ， $W^{YX}\mathbf{g}^X(\mathbf{x})$ ， $W^{YY}\mathbf{g}^Y(\mathbf{y})$ をそれぞれ $\widetilde{W}^{XX}(\mathbf{g}^X(\mathbf{x}))$ ，

$\widetilde{W}^{XY}(\mathbf{g}^Y(\mathbf{y}))$, $\widetilde{W}^{YX}(\mathbf{g}^X(\mathbf{x}))$, $\widetilde{W}^{YY}(\mathbf{g}^Y(\mathbf{y}))$ と置く．新しい $\widetilde{W}^{XX}(\cdot)$, $\widetilde{W}^{XY}(\cdot)$, $\widetilde{W}^{YX}(\cdot)$, $\widetilde{W}^{YY}(\cdot)$ は行列ではなくベクトル関数であり，具体的には次節で定義する．

すなわち，式 (4.5),(4.6) はそれぞれ以下ようになる．

$$\tau \frac{d\mathbf{x}}{dt} = -\mathbf{x} + \widetilde{W}^{XX}(\mathbf{g}^X(\mathbf{x})) + \widetilde{W}^{XY}(\mathbf{g}^Y(\mathbf{y})) + \mathbf{b}^X \quad (4.8)$$

$$\tau \frac{d\mathbf{y}}{dt} = -\mathbf{y} + \widetilde{W}^{YX}(\mathbf{g}^X(\mathbf{x})) + \widetilde{W}^{YY}(\mathbf{g}^Y(\mathbf{y})) + \mathbf{b}^Y \quad (4.9)$$

本章では，これらのベクトル関数を結合荷重関数と呼ぶ．

4.3.2 可視空間から隠れ空間への写像

隠れ素子を導入した相互結合型ネットワークでは，可視空間と状態空間における力学系の性質の関係から大きく2種に分けることができる．

- (1) 可視空間の力学的挙動が元の力学系の挙動から大きく変わってしまう場合．たとえば部分観測によって，元の空間では起こらない軌道の交差が可視空間で起こる場合などがこれに相当する．
- (2) 可視空間の力学的挙動が元の力学系の挙動を十分に反映している場合．すなわち射影によって，力学的性質があまり失われない場合．

(1) の場合は，連想過程に一種の文脈依存性が生じ，脳での連想過程を考える上でも興味深い設定である．しかし本研究では，隠れ素子は可視素子の能力の向上のための補助素子として位置付ける．すなわち，可視空間での表現を，高次元の状態空間での表現に変換することによって，想起過程を改良することを目的としている．たとえば，可視空間だけを用いた場合に，偽記憶の引き込み領域が真の記憶のそれよりも大きく想起が失敗しやすい時にも，高次元の想起過程を用いることによって改善できる可能性がある．この目的のために，本章では上記 (2) の立場を取る．すなわち，可視空間と状態空間とで力学的挙動があまり変わらないようにする．

可視変数 x と状態変数 u とが一对一の関係にあれば上記 (2) の性質を実現できる。すなわち、可視空間から隠れ空間への写像を h とすれば、隠れ変数 y は次のようにあらわせる。

$$\mathbf{y}(t) = \mathbf{h}(\mathbf{x}(t)) \quad (t \geq 0) \quad (4.10)$$

h として線形写像を用いることにする。

$$\mathbf{y}(t) = A\mathbf{x}(t) + \boldsymbol{\theta} \quad (4.11)$$

A は $R \times N$ 行列、 $\boldsymbol{\theta}$ は R 次ベクトルである。連立方程式 (4.8),(4.9) が式 (4.11) の解を持つためには、以下の条件が成り立てば十分である。

$$\widetilde{W}^{YX}(\mathbf{g}^X(\mathbf{x})) = A\widetilde{W}^{XX}(\mathbf{g}^X(\mathbf{x})) \quad (4.12)$$

$$\widetilde{W}^{YY}(\mathbf{g}^Y(A\mathbf{x} + \boldsymbol{\theta})) = A\widetilde{W}^{XY}(\mathbf{g}^Y(A\mathbf{x} + \boldsymbol{\theta})) \quad (4.13)$$

$$\mathbf{b}^Y = A\mathbf{b}^X + \boldsymbol{\theta} \quad (4.14)$$

$$\mathbf{y}(0) = A\mathbf{x}(0) + \boldsymbol{\theta} \quad (4.15)$$

この条件のもとで連立方程式 (4.8),(4.9) は以下の一つの方程式になる。

$$\begin{aligned} \tau \frac{d\mathbf{x}}{dt} = & -\mathbf{x} + \widetilde{W}^{XX}(\mathbf{g}^X(\mathbf{x})) \\ & + \widetilde{W}^{XY}(\mathbf{g}^Y(A\mathbf{x} + \boldsymbol{\theta})) + \mathbf{b}^X \end{aligned} \quad (4.16)$$

式 (4.11) により状態変数 u の有効次元は可視変数 x の次元 N に下がっているので、可視空間の力学系挙動は元の力学系挙動を完全に反映している。

可視部と隠れ部が線形の関係にあるとする式 (4.11) は Kimura ら [30] によるものである。本章ではさらに、式 (4.8),(4.9) のように結合荷重行列を結合荷重関数と置き換えることにより、4.4 節で述べるハイブリッド型学習を可能にするものである。

4.4. 連想記憶モデルの構築

本章では、式 (4.16) であらわされる隠れ素子付きネットワークを連想記憶モデルとして適用する手法について述べる。

本章では式 (4.16) の右辺第二項, 第三項を, それぞれ次のように置く.

$$\widetilde{W}^{XX}(\mathbf{g}^X(\mathbf{x})) \equiv \frac{a}{2} \mathbf{F}^h[J\mathbf{g}^X(\mathbf{x})] \quad (4.17)$$

$$\begin{aligned} \widetilde{W}^{XY}(\mathbf{g}^Y(A\mathbf{x} + \boldsymbol{\theta})) \\ \equiv \frac{2-a}{2} \mathbf{F}^b[B\{\mathbf{g}^Y(A\mathbf{x} + \boldsymbol{\theta}) + \boldsymbol{\phi}\}] \end{aligned} \quad (4.18)$$

$\mathbf{F}^h[\cdot], \mathbf{F}^b[\cdot]$ はベクトル関数である. J は $N \times N$ 行列, B は $N \times R$ 行列, $\boldsymbol{\phi}$ は R 次のベクトル, a は定数である. $\boldsymbol{\theta}$ も含め, これらは何らかの手段によって決まるとする (4.4.1 項の学習法から決まる).

式 (4.12), (4.13) の条件から $\widetilde{W}^{YX}(\mathbf{g}^X(\mathbf{x}))$, $\widetilde{W}^{YY}(\mathbf{g}^Y(A\mathbf{x} + \boldsymbol{\theta}))$ は以下で与えられる.

$$\widetilde{W}^{YX}(\mathbf{g}^X(\mathbf{x})) \equiv A \frac{a}{2} \mathbf{F}^h[J\mathbf{g}^X(\mathbf{x})] \quad (4.19)$$

$$\begin{aligned} \widetilde{W}^{YY}(\mathbf{g}^Y(A\mathbf{x} + \boldsymbol{\theta})) \\ \equiv A \frac{2-a}{2} \mathbf{F}^b[B\{\mathbf{g}^Y(A\mathbf{x} + \boldsymbol{\theta}) + \boldsymbol{\phi}\}] \end{aligned} \quad (4.20)$$

式 (4.14) については, $\mathbf{b}^X = \mathbf{0}$ と置くことで,

$$\mathbf{b}^Y = \boldsymbol{\theta} \quad (4.21)$$

とする.

出力関数 $g_i^X(x_i)$, $g_j^Y(y_j)$ には次のような双曲正接を用いる [26].

$$g_i^X(x_i) = \tanh(10x_i) \quad (1 \leq i \leq N) \quad (4.22)$$

$$g_j^Y(y_j) = \tanh\left(\frac{1}{10}y_j\right) \quad (1 \leq j \leq R) \quad (4.23)$$

式 (4.22) の双曲正接における逆温度パラメータを 10 と大きくしたのは, 出力に関わる可視素子が ± 1 に近い値をとりやすくしたためであり, 一方で, 式 (4.23) の逆温度パラメータを $1/10$ としたのは, 隠れ素子が中間的な値をとりやすいようにしたためである.

さらに、ベクトル関数 $F^h[\cdot], F^b[\cdot]$ のそれぞれの要素関数 $F_i^h[\cdot], F_i^b[\cdot]$ にも、式 (4.22), (4.23) と同様に双曲正接を用いる。

$$F_i^h[s] = \tanh(50s) \quad (4.24)$$

$$F_i^b[s] = \tanh(10s) \quad (4.25)$$

結果として、式 (4.16) は、

$$\begin{aligned} \tau \frac{d\mathbf{x}}{dt} = & -\mathbf{x} + \frac{a}{2} \mathbf{F}^h[J\mathbf{g}^X(\mathbf{x})] \\ & + \frac{2-a}{2} \mathbf{F}^b[B\{\mathbf{g}^Y(A\mathbf{x} + \boldsymbol{\theta}) + \boldsymbol{\phi}\}] \end{aligned} \quad (4.26)$$

となる。 a は第二項と第三項の影響力の大きさを調節する定数である。

4.4.1 学習法

本章ではモデルに記憶させるパターン(以下記憶パターン)に、各要素が $1/2$ の確率で ± 1 の離散値をとる P 個の N 次元ベクトル $\xi^\mu, (1 \leq \mu \leq P)$ を用いることにする。 μ 番目の記憶パターンの i 番目の要素を ξ_i^μ と置く。

J は次に示すヘップ則により決める。

$$J_{ij} = \frac{1}{N} \sum_{\mu=1}^P \xi_i^\mu \xi_j^\mu \quad (i \neq j) \quad (4.27)$$

ただし、 $J_{ii} = 0$ である。

図 4.3 に示すように式 (4.26) の第三項は 2 層の MLP として解釈できる。 A は入力層から中間層への結合荷重行列、 $\boldsymbol{\theta}$ は中間層での閾値、 B は中間層から出力層への結合荷重行列、 $B\boldsymbol{\phi}$ は出力層の閾値となる。このとき、入力層に \mathbf{x} を入力したときの中間層の状態変数は $A\mathbf{x} + \boldsymbol{\theta}$ 、出力状態は $\mathbf{g}^Y(A\mathbf{x} + \boldsymbol{\theta})$ 、出力層の状態変数は $B\{\mathbf{g}^Y(A\mathbf{x} + \boldsymbol{\theta}) + \boldsymbol{\phi}\}$ となる。この 2 層の MLP のパラメータを誤差逆伝播法 [38] により決定する。

学習サンプルペアには入力、出力ともに記憶パターン $\xi^\mu, (1 \leq \mu \leq P)$ を用いる。つまり、 ξ^μ が入力されれば ξ^μ を出力するように学習させる。

本章では学習サンプルペアの入力と出力が同一であり，それらには記憶パターンのみを用いているので，本章の誤差逆伝播法は式 (4.27) と同様に教師なし学習と見なせる．したがって，使用した学習法はハイブリッド型の教師なし学習である．

パラメータ a はハイブリッドした二つのモジュールの重みを変えるもので，その決定は重要である．本章では以下のような実験により， a を決めた．可視素子数 $N = 200$, 隠れ素子数 $R = 30$, 記憶パターン数 $P = 10$ として学習を行う．学習後のネットワークに 20 個のランダムパターンを入力し，偽記憶へ収束した回数をカウントする． a に対する偽記憶への収束回数をプロットしたものを図 4.4 に示す．

図 4.4 から $0.9390 \leq a \leq 0.9400$ において偽記憶への収束回数が最小になるので，その中間値である $a = 0.9395$ を a の値として用いた． $a = 1$ の時バランスされているので， a が 1 より小さいということは，式 (4.26) の第三項 (MLP) の影響力を幾分大きくしていることになる．

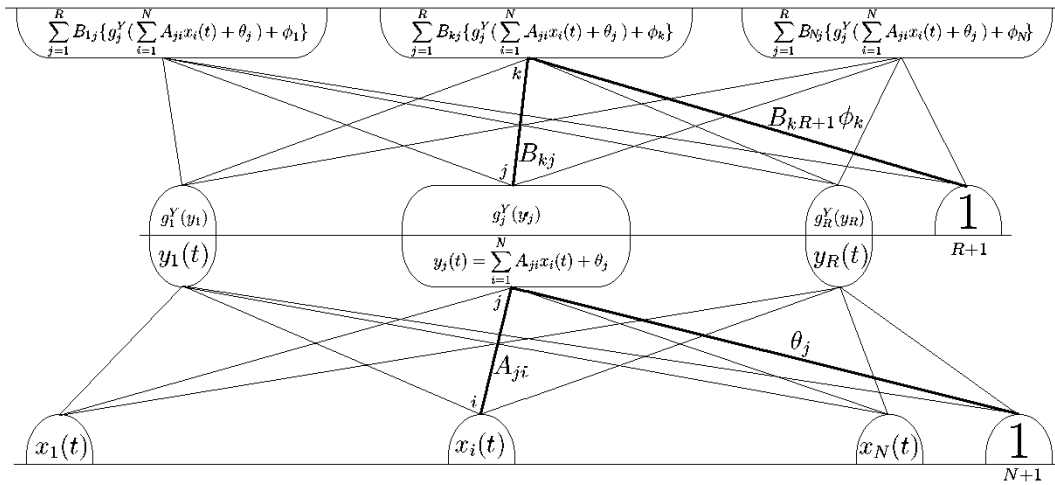


図 4.3 式 (4.26) の第三項 [] 内 $B\{g^Y(Ax + \theta) + \phi\}$ の展開図．入力層の i 番目の素子と中間層の j 番目の素子の結合荷重は A_{ij} ，中間層の j 番目の素子と出力層の k 番目の素子の結合荷重は B_{kj} ，中間層の j 番目の素子にかかるバイアスは θ_j ，出力層の k 番目の素子にかかるバイアスは $B_{kR+1}\phi_k$ となる．これらの結合線は分かりやすいように他の線と区別して太線を用いている．

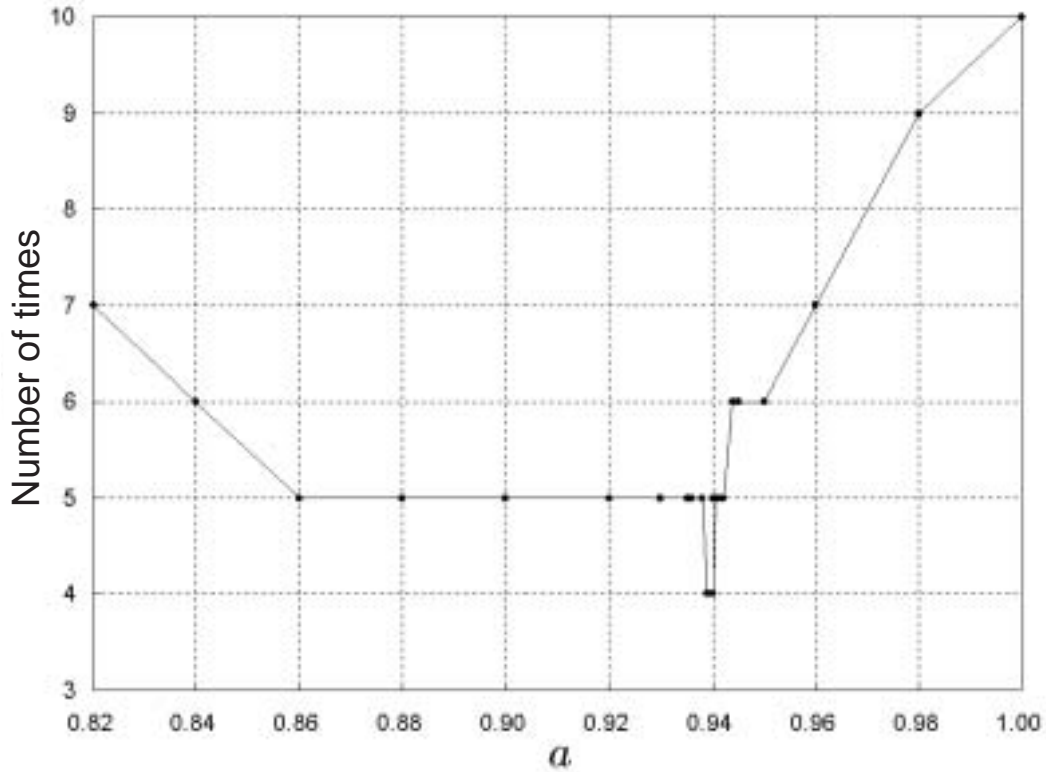


図 4.4 横軸に α , 縦軸に偽記憶への収束回数をプロットした図 .

4.4.2 離散型

式 (4.22), (4.23), (4.24), (4.25) の双曲正接の傾きを無限大と見なすことにより , 式 (4.26) を次の離散型の式に置き換えることができる [26] .

$$\begin{aligned}
 S_{(T+1)} = & \frac{\alpha}{2} \operatorname{sgn}[JS_{(T)}] \\
 & + \frac{2-\alpha}{2} \operatorname{sgn}[B\{g^Y(AS_{(T)} + \theta) + \phi\}] \quad (4.28)
 \end{aligned}$$

ここで , $S_{(T)}$ は離散時刻 T における可視部の状態をあらわす . また , α は第一項と第二項の影響力の大きさを調節する定数であり , 本章では図 4.4 の実験結果にならない , 1 よりも幾分小さい値として $\alpha = 0.9999$ とした . したがって , S の各要素 S_i は ± 1 または ± 0.0001 の離散値をとる .

区別のために，式 (4.26) を用いたモデルを連続型，式 (4.28) を用いたモデルを離散型と呼ぶことにする．

4.5. シミュレーション

本章で行うシミュレーションはすべて，可視素子数 $N = 200$ ，隠れ素子数 $R = 30$ とした．主に $P = 10$ としてシミュレーションを行うので，誤差逆伝播法によって 10 個すべての記憶パターンを学習可能するのに十分な中間層の素子数となるように決めた．すなわち $R = 30$ であれば，誤差逆伝播法による学習後の階層型ネットワークに，10 個の記憶パターンを入力してもすべて正しく出力される．

4.5.1 偽記憶への収束率

本項では，任意のパターンを入力した場合に提案したモデルがどのようなパターンに収束するかを調べる．シミュレーションの手順を以下に示す．

手順 1 任意の入力パターンを状態更新式 (連続型の場合は式 (4.26)，離散型の場合は式 (4.28)) に入力．

手順 2 ネットワークを収束させ出力パターンを得る．

手順 3 出力パターンと一番近い記憶パターン ξ^μ とのオーバーラップ m^μ を記録．

手順 4 手順 1 から 3 を 1000 回繰り返す．

ここで，連続型用の入力パターンの要素 $x_i(0)$ は閉区間 $[-1, +1]$ の間で一様なランダム値をとらせることにより生成し，出力パターンは式 (4.26) の収束状態 $x(\infty)$ を用いて $g^X(x(\infty))$ とする．離散型用の入力パターンの要素は ± 1 の離散値を $1/2$ の確率でランダムに割り振ることにより生成し，出力パターンは式 (4.28) の収束状態 $\text{sgn}(S(\infty))$ とする．出力パターン X と記憶パターン ξ^μ とのオーバーラップ m^μ は以下により定義する．

$$m^\mu = \frac{1}{N} X \cdot \xi^\mu \quad (4.29)$$

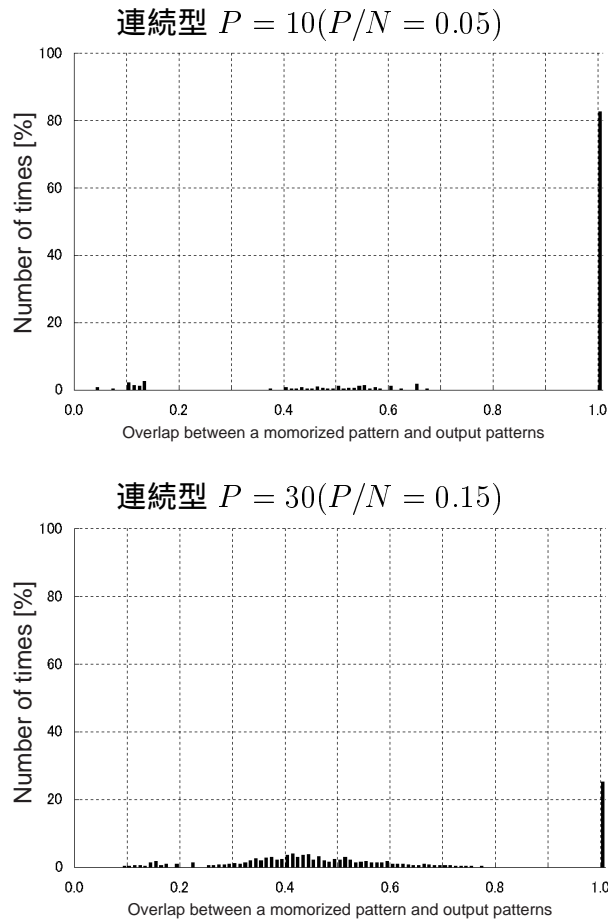


図 4.5 連続型における出力パターンと記憶パターンのオーバーラップをあらわすグラフ。

ここで、 \cdot は二つのベクトルの内積を意味する。 m^μ は -1 から 1 の間の値をとり、 $m^\mu = 1$ の場合は出力パターン X と記憶パターン ξ^μ が完全に一致することを示す。逆に $m^\mu = 0$ の場合は $1/2$ の素子が一致し、双方のパターンの相関がゼロであることを示す。

連続型と離散型におけるシミュレーション結果をそれぞれ図 4.5, 図 4.6 に示す。一番近い記憶パターンとのオーバーラップであるので、 0.0 未満のオーバーラップに収束することはない。横軸に収束パターンとそれに最も近い記憶パターンとのオーバーラップ、縦軸に収束回数の割合としている。いずれの場合もネットワークは

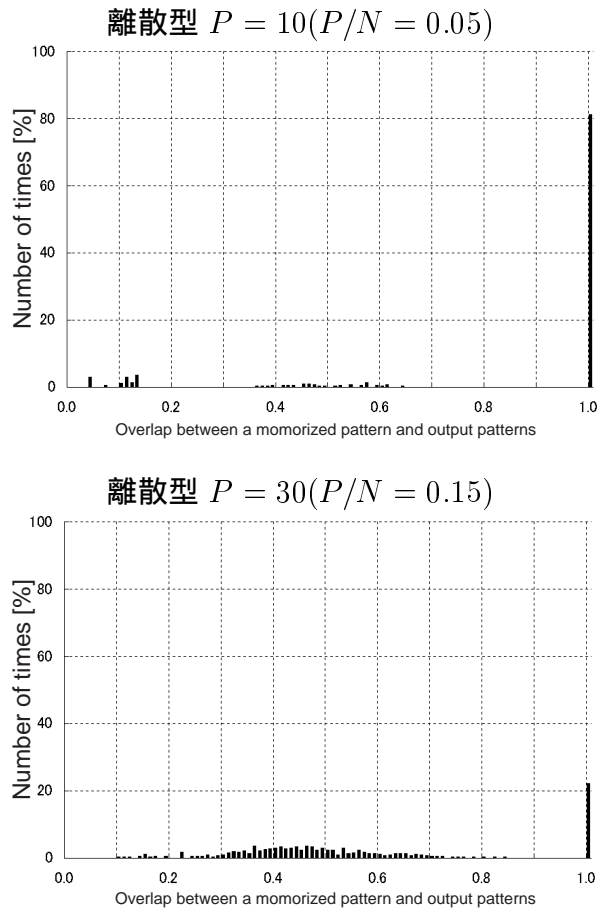


図 4.6 離散型における出力パターンと記憶パターンのオーバーラップをあらわすグラフ。

すべて収束した。オーバーラップ 1.0 以外はすべて偽記憶をあらわしている。 $P = 10$ の場合、連続型は約 82%，離散型は図 4.6 から約 81% が記憶パターンに収束した。

4.5.2 臨界オーバーラップ

本項では、提案したモデルの臨界オーバーラップを求める。臨界オーバーラップとは、ノイズを加えた記憶パターンを入力パターンとして用いた場合に、元の記憶パターンを出力することができる、記憶パターンと入力パターンの最小オーバーラップである。シミュレーションの手順を以下に示す。

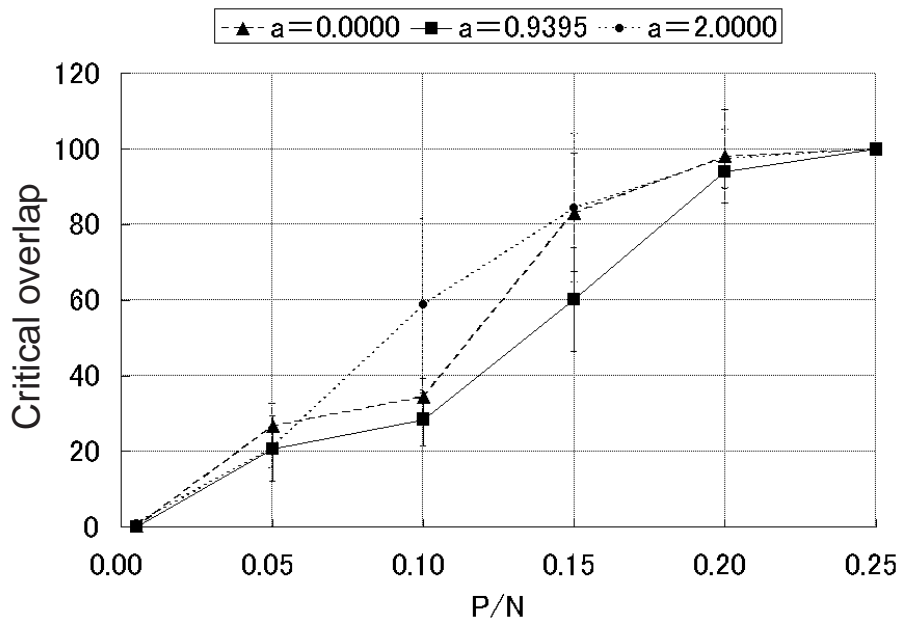


図 4.7 臨界オーバーラップ．連続型．

手順 1 記憶パターンにノイズを加え，入力パターンを作成．

手順 2 入力パターンを状態更新式 (連続型の場合は式 (4.26)，離散型の場合は式 (4.28)) に入力．

手順 3 ネットワークを収束させ出力パターンを得る．

手順 4 元の記憶パターンに収束した場合は入力パターンにさらなるノイズを加えることにより，入力パターンと記憶パターンのオーバーラップを減らし，手順 2 に戻る．

手順 5 元の記憶パターンに収束可能な最小のオーバーラップを臨界オーバーラップとして求める．

手順 6 新しい記憶パターンを用意し，手順 1 から 5 を 50 回繰り返し，臨界オーバーラップの平均を求める．

図 4.7 に連続型の場合の臨界オーバーラップ，図 4.8 に離散型の場合の臨界オーバーラップを求めた結果をそれぞれ示す．両グラフとも横軸が記憶率 (P/N)，縦軸

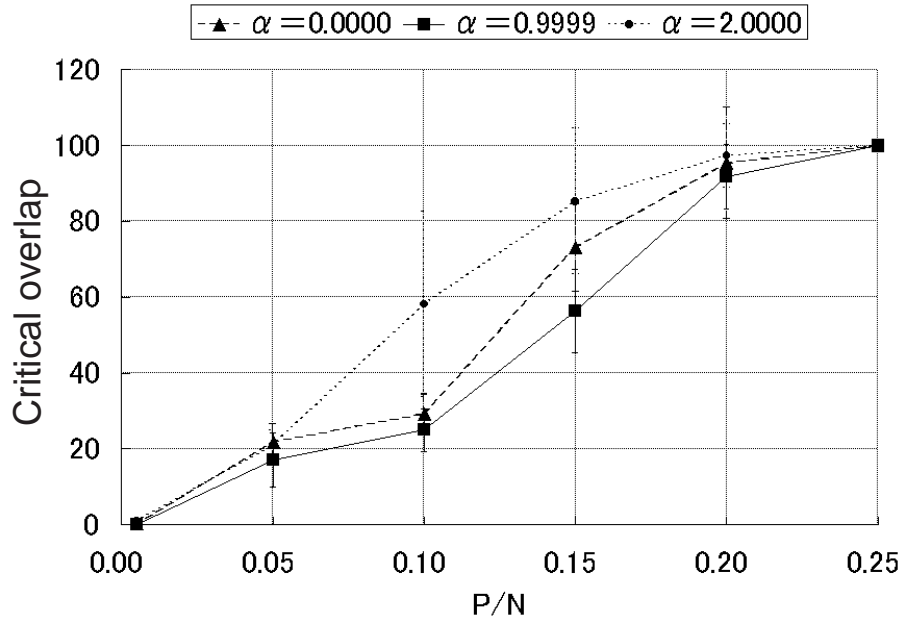


図 4.8 臨界オーバーラップ．離散型．

が臨界オーバーラップをあらわしている．比較のために，隠れ素子が付かない場合 ($a = \alpha = 0.0000$) と階層型ネットワークだけの場合 ($a = \alpha = 2.0000$) のグラフを載せている．いずれの場合も提案したモデルが最も小さい臨界オーバーラップを持つことが分かる．小さい臨界オーバーラップを持つことは広い引き込み領域を持つことと等価であるので，提案モデルは広い引き込み領域を持つことになる．

4.6. 本章のむすび

本章では，隠れ素子を可視素子の能力向上のための補助素子として用いた．可視空間から隠れ空間への写像には線形写像を用いた．結合荷重行列を結合荷重関数と置くことにより，連想記憶モデルへの適用を可能にした．そして，提案モデルが広い引き込み領域を持つことにより，偽記憶へ収束しにくいことをシミュレーションにより示した．

4.7. 付録：誤差逆伝播法

本節ではハイブリッド型学習の一つとして用いた誤差逆伝播法 [38] についての解説を行う。

まずネットワークの定式化を行う。階層型ネットワークの概念図を図 4.9 に示

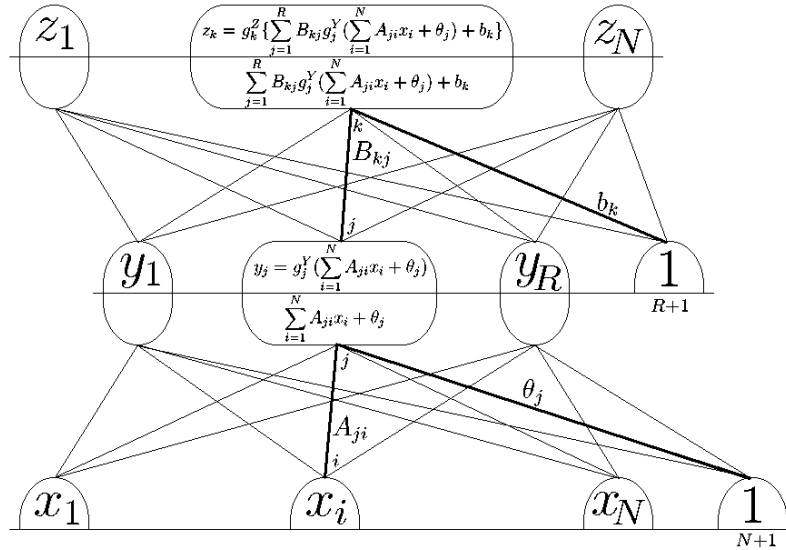


図 4.9 階層型ネットワークの概念図。

す。入力層の状態を入力状態ベクトル x により表現する。このベクトルの次元は入力層の素子の数 N に等しい。同様に中間層の出力状態ベクトルを y 、出力層の出力状態ベクトルを z とする。中間層の素子数を R 、出力層の素子数は入力層と同じ N とする。一般には、出力層の素子数は入力層のそれとは異なるが、本章で述べたハイブリッド型連想記憶モデルでは入力層の素子数と出力層の素子数が等しい場合をとり扱ったので、本節でもその場合を考える。ここで、入力層の i 番目の素子と中間層の j 番目の素子を結ぶ結合荷重を A_{ji} とすれば、中間層の出力状態 y_j は次のようになる。

$$y_j = g_j^Y \left(\sum_{i=1}^N A_{ji} x_i + \theta_j \right) \quad (4.30)$$

入力層に入力状態値が常に 1 である素子を 1 つ付け加えることにより、その素子

と中間層の j 番目の素子を結ぶ結合荷重を θ_j とした． $g_j^Y(s)$ は中間層における j 番目の素子の出力関数である．

同様に中間層の j 番目の素子と出力層の k 番目を結ぶ結合荷重を B_{kj} とすれば，中間層の出力状態 z_k は次のようになる．

$$z_k = g_k^Z \left(\sum_{j=1}^R B_{kj} y_j + b_k \right) \quad (4.31)$$

$$= g_k^Z \left\{ \sum_{j=1}^R B_{kj} g_j^Y \left(\sum_{i=1}^N A_{ji} x_i + \theta_j \right) + b_k \right\} \quad (4.32)$$

ここでも中間層に出力状態値が常に 1 である素子を 1 つ付け加えることにより，その素子と出力層の k 番目の素子を結ぶ結合荷重を b_k とした．ただし，中間層に付け加えた素子は出力層の素子とのみ結合を持ち，入力層の素子との結合は持たない． $g_k^Z(s)$ は出力層における k 番目の素子の出力関数である．誤差逆伝播法を用いる場合，出力関数 $g_j^Y(s), g_k^Z(s)$ はともに微分可能でなければならないので，通常は双曲正接や線形関数が用いられる．

次に結合荷重を決めるために学習を行う．学習を行うにはまず学習サンプルペアを決めなくてはならない． μ 番目の入力サンプルを ζ^μ ，教師サンプルを $\bar{\zeta}^\mu$ とする． μ 番目の入力サンプルを入力した場合の中間層における実際の出力を

$$y_j^\mu = g_j^Y \left(\sum_{i=1}^N A_{ji} \zeta_i^\mu + \theta_j \right) \quad (4.33)$$

とする．同様に出力層における実際の出力を

$$z_k^\mu = g_k^Z \left(\sum_{j=1}^R B_{kj} y_j^\mu + b_k \right) \quad (4.34)$$

$$= g_k^Z \left\{ \sum_{j=1}^R B_{kj} g_j^Y \left(\sum_{i=1}^N A_{ji} \zeta_i^\mu + \theta_j \right) + b_k \right\} \quad (4.35)$$

とする．ここで，学習サンプルペアの数を Q とすれば学習の精度を評価する学習誤差 E を次のように定義できる．

$$E = \frac{1}{2} \sum_{\mu=1}^Q \lambda(\mu) \sum_{k=1}^N (z_k^\mu - \bar{\zeta}_k^\mu)^2 \quad (4.36)$$

ここで， $\lambda(\mu)$ はそれぞれの学習サンプルペアにおける重要さの度をあらわす関数である．誤差逆伝播法の基本的な考え方は，学習誤差 E が最小になるように，勾配法を用いて結合荷重を修正することである． $E = 0$ であれば，提示された学習サンプルペアに対して学習が完全に達成されたことになる．

式 (4.36) の E は，入力サンプル $\bar{\zeta}_k^\mu$ を通じて全ての結合荷重 $\{A_{ji}, B_{kj}\}$ の関数になっている．したがって， E の B_{kj} での微係数に小さな負の数を掛け量だけ B_{kj} を変化させると E が減少する． ϵ を微小な正の定数として，この事実を式であらわすと次のようになる．

$$\Delta B_{kj} = -\epsilon \frac{\partial E}{\partial B_{kj}} \quad (4.37)$$

与えられた B_{kj} に対して E の微係数が正であれば， B_{kj} を少し減らすことにより E は減少し，微係数が負であれば B_{kj} を増加させる．

学習誤差 (4.36) に対して式 (4.37) を適用することにより，

$$\Delta B_{kj} = -\epsilon \frac{\partial E}{\partial B_{kj}} \quad (4.38)$$

$$= \epsilon \sum_{\mu=1}^Q \lambda(\mu) \sum_{k=1}^N (\bar{\zeta}_k^\mu - z_k^\mu) g_k^{\prime Z} \left(\sum_{l=1}^R B_{kl} y_l^\mu + b_k \right) y_j^\mu \quad (4.39)$$

$$= \epsilon \sum_{\mu=1}^Q \lambda(\mu) \delta_k^\mu y_j^\mu \quad (4.40)$$

となる．ここで $g_j^{\prime Y}(s)$ は $g_j^Y(s)$ の導関数であり，

$$\delta_k^\mu = \sum_{k=1}^N (\bar{\zeta}_k^\mu - z_k^\mu) g_k^{\prime Z} \left(\sum_{l=1}^R B_{kl} y_l^\mu + b_k \right) \quad (4.41)$$

とおいた． $\bar{\zeta}_k^\mu$ と z_k^μ の差はネットワークの出力誤差であるので， δ_k^μ はこの誤差に關係する量である．

中間層と入力層の結合 A_{ji} の修正も同様にして求めることができる．学習誤差 E がどう A_{ji} に依存しているかは，式 (4.35) より明らかである．これを用いれば，

$$\Delta A_{ji} = -\epsilon \frac{\partial E}{\partial A_{ji}} \quad (4.42)$$

$$= -\epsilon \sum_{k=1}^N \sum_{\mu=1}^Q \frac{\partial E}{\partial y_j^\mu} \frac{\partial y_j^\mu}{\partial A_{ji}} \quad (4.43)$$

$$= \epsilon \sum_{\mu=1}^Q \lambda(\mu) g_j^Y \left(\sum_{i=1}^N A_{ji} \zeta_i^\mu + \theta_j \right) \sum_{k=1}^N B_{kj} \delta_k^\mu \zeta_i^\mu \quad (4.44)$$

$$= \epsilon \sum_{\mu=1}^Q \lambda(\mu) \gamma_j^\mu \zeta_i^\mu \quad (4.45)$$

となる．ここで $g_j^Y(s)$ は $g_j^Y(s)$ の導関数であり，

$$\gamma_j^\mu = g_j^Y \left(\sum_{i=1}^N A_{ji} \zeta_i^\mu + \theta_j \right) \sum_{k=1}^N B_{kj} \delta_k^\mu \quad (4.46)$$

とおいた．式 (4.40) と式 (4.45) は，誤差に関する量 δ_k^μ の情報をもとにして結合荷重を修正する規則を与えている． δ_k^μ は出力層での誤差， γ_j^μ は中間層での誤差によって引き起こされる量であるので，式 (4.46) は誤差が出力層から中間層に向かって伝えられることをあらわしている．つまり修正情報の伝わる方向は，信号の伝搬方向とは逆である．この学習法を誤差逆伝播法と呼ぶ．

誤差逆伝播法で注意すべきことは，誤差計算が前方向への情報の流れとよく類似している点である．すなわち，ある素子の学習に使われている情報は，後の素子から得られる情報のみであり，学習の局所性が保たれていることになる．これは，人工的な神経回路計算機（局所結合のみを持つマルチプロセッサシステム）をハードウェア化するとき学習則に要求される性質である．ただし，実際の生体の神経回路においては， δ_k^μ といった値が神経軸策を通して逆向きに伝わるはずはなく，誤差逆伝播法はこういった意味から，実際の脳における生理学的学習モデルにはなりえないことになる．

第5章 総括

本論文では，情報復元についての統計力学的な立場から議論した．

第1章で本論文の構成と位置づけを述べ，第2章ではこれまでの先行研究を紹介した．まず，2.1節では符号化しない場合の情報復元として画像修復問題を取り上げ，ガウス模型による画像修復の解説を行った．フーリエ変換を用いることで修復結果が簡単に求まることを示した．また，画像修復問題が有効ハミルトニアン基底状態探索問題に帰着できることを紹介した．

2.2節では Sourlas 符号に着目することで，冗長化することで符号化した場合の情報復元について議論した．画像修復問題と同様に，Sourlas 符号を用いた誤り訂正符号問題が，ハミルトニアン基底状態探索問題に置き換えることができることを見てきた．特に2体相互作用の場合は，冗長化された送信信号 J^{in} が，一つのパターンを記憶した場合の連想記憶モデルの学習則と同一のものとなることを示した．

2.3節では Sourlas 符号の拡張モデルとして連想記憶モデルについて述べた．Sourlas 符号と同様に Hopfield 型の連想記憶モデルのハミルトニアンを定義することで想起過程がハミルトニアンの緩和過程に相当することを示した．特に，2体相互作用の Sourlas 符号のハミルトニアンと一つのパターンを記憶した場合の連想記憶モデルのハミルトニアンが等しいものになっていることを示した．さらにシミュレーションを用いることで連想記憶モデルの主な性質を見てきた．その中で偽記憶への収束が連想記憶モデルの問題の一つになっていることを紹介した．第2章で先行研究として述べたモデルはいずれもハミルトニアン基底状態探索問題に置き換えられることを示した．

第3章では，さらに画像修復問題について統計力学的な解析を行うことで，その定性的な性質を求めた．特に劣化画像と修復画像の平均二乗誤差の厳密な解や，

フーリエ変換を用いた場合の周辺尤度などを導出してた。また、真のノイズ通信路と実際に用いるモデルが異なる場合に注目することで、その場合の解析を行った。特に、真のモデルとして空間的な相関を持つガウスノイズを用い、実際に用いるモデルには空間的な相関のないガウスノイズを用いた。この場合における修復誤差最小化基準による解析的な修復誤差を求めたあと、周辺尤度最大化基準による数値実験的な修復誤差を求めた。数値実験では、画像を修復するのに必要な、ハイパーパラメータを求めるために周辺尤度最大化を用いていたが、二つのモデルが異なる場合はハイパーパラメータがうまく推定できないことを示した。そのため、真のモデルと実際に用いるモデルが異なる場合において周辺尤度最大化基準を用いたときは、修復誤差最小化基準で解析的に得られた結果ほど修復誤差を小さくすることはできず、最高の性能が発揮できないことが分かった。さらに、人工画像と自然画像の両方に対してコンピュータシミュレーションを行っており、その結果は解析結果を指示していることを確認した。

第4章では2.3節で紹介した連想記憶モデルの問題の一つである、偽記憶に収束しにくいモデルとして、ハイブリッド型の連想記憶モデルを提案した。提案モデルでは、連想処理専用の隠れ部を付け加えた。可視部と隠れ部の写像に線形写像を用いることにより、可視部 隠れ部 可視部の一連の再帰的結合を、二層の階層型パーセプトロンと見なし、可視部と隠れ部の結合の学習には誤差逆伝播法を、そして可視部の相互結合にはヘップ則を用いることで、ハイブリッド型の学習を実現した。シミュレーションにより、提案モデルが広い引き込み領域を持つことにより、偽記憶へ収束しにくいことを示した。

以上のように本論文では、符号化しなかった場合の情報復元として画像修復問題に着目し、また符号化した場合の情報復元については Sourlas 符号と Hopfield 型の連想記憶モデルについて着目することで議論した。いずれの場合もハミルトニアンを定義することにより、基底状態探索問題に置き換えることができることを見てきた。また、統計力学的な手法がモデルの定性的な解析に有利であることを理解してきた。

第3章で提案した空間的な相関を持つノイズモデルには、論文中で紹介したように特異点があることが分かっている。この特異点を解消すれば、工学的に価値

のある実用的な応用も可能になる．特異点の解消法としては，まずノイズモデルのカーネルを簡単にすることがあげられる．第3章で用いたノイズモデルのカーネルは統計的な解析を行うために，やや複雑な形をしている．これを簡単にすることで特異点を解消できる可能性がある．また，修復に使うカーネルそのものを代えることも可能である．第3章では，修復にモデルが異なる場合，周辺尤度最大化によるハイパーパラメータ推定が困難であることを述べたが，モデルが異なる場合であっても，元のモデルのカーネルと相性の良いカーネルが存在する可能性がある．したがって，異なるモデルであっても，特異点を持たず，元のモデルのカーネルと相性の良いカーネルを探し出すことができれば，修復を旨く行かせることが可能になるであろう．

謝辞

本論文を執筆するにあたり，奈良先端科学技術大学院大学情報科学研究科 伊藤実教授には御指導と御教授を賜りました．ここで伊藤教授に謹んで感謝申し上げます．伊藤教授には，筆者が本論文の執筆を開始する際に，指針や御助言を与えてくださいました．さらには，本論文の校正および主査をして頂きました．重ねて深謝申し上げます．

また，本論文の第4章の内容を研究するにあたり，奈良先端科学技術大学院大学情報科学研究科石井信教授には研究の具体的な指針，まとめ方，その他多数の御助言を賜りました．石井信教授に謹んで深謝申し上げます．本論文の第4章の研究が巧く行ったのは，石井教授の貴重な御教授に基づくものであります．重ねて深謝申し上げます．

本論文の第3章の内容は著者が特別派遣学生として理化学研究所において研修生をさせて頂いた期間に行われた研究であります．理化学研究所脳科学総合研究センター脳数理研究チーム副チームリーダーの岡田真人先生には多大なる御指導，御教授を頂きました．ここに岡田先生に深く感謝申し上げます．その他多数の事においてもお世話になりました．改めて感謝申し上げます．

本論文の第4章の内容は著者が九州大学大学院修士課程時代に開始し，奈良先端科学技術大学院大学の博士後期課程一年において完成した研究であります．九州大学大学院システム情報科学研究科 高橋規一助教授には研究の機会を与えて頂くと共に，修士時代において御指導頂きました．ここに厚く御礼申し上げます．

理化学研究所脳科学総合研究センター脳数理研究チームチームリーダーの甘利俊一教授（東京大学名誉教授）は著者が理化学研究所研修生として理研で研究させて頂くことを許可して下さり，大変恵まれた環境において研究させて頂く機会を著者に与えてくださいました．ここに深く感謝の意を申し上げさせて頂きます．

奈良先端科学技術大学院大学情報科学研究科 作村勇一助手には，研究に関し，多くの御助言を頂きました．ここに厚く感謝申し上げます．東北大学大学院情報科学研究科 田中和之助教授には第3章のに関し御助言を賜りました．ここに厚く御礼申し上げます．

研究活動全般にわたってサポートして頂いた同じ研究部屋の理化学研究所 JRA

の西村信一医師（東京大学大学院医学研究科）と理化学研究所研修生の和田浩司氏（埼玉大学大学院理工学研究科）ならびに理化学研究所 JRA の井上真郷医師（京都大学大学院医学研究科）には日頃から大変お世話になりました．感謝の極みにございます．

最後に，何かとお世話くださった，理化学研究所，奈良先端科学技術大学院大学，九州大学大学院の皆様にご心より御礼申し上げます．

参考文献

- [1] D. Geman, “Random Fields and Inverse Problems in Imaging,” Lecture Notes in Mathematics, no.1427, pp.113-193, Springer-Verlag, New York, 1990.
- [2] R. Chellappa and A. Jain (eds), “Markov Random Fields: Theory and Applications,” Academic Press, New York, 1993.
- [3] S. Z. Li, “Markov Random Field Modeling in Computer Vision,” Springer-Verlag, Tokyo, 1995.
- [4] J. Besag, “On the statistical analysis of dirty pictures,” J. Royal Statistical Society B (with discussion), vol.48, no.3, pp.259-302, 1986.
- [5] J. Marroquin, S. Mitter and T. Poggio, “Probabilistic solution of ill-posed problems in computational vision,” Journal of the American Statistical Association, vol.82, no.397, no.3, pp.76-89, 1987.
- [6] “間瀬滋, 武田純, 空間データモデリング”, 共立出版, 2001.
- [7] J. M. Pryce and A. D. Bruce, “Statistical mechanics of image restoration,” Journal of Physics A, vol.28, no.3, pp.511-532, 1995.
- [8] 西森秀稔, “スピングラス理論と情報統計物理学,” 新物理学選書, 岩波書店, 1999.
- [9] 田中和之, “統計力学的手法をもとにした画像修復,” 日本物理学会誌, vol.5’, no.1, pp.25-ee, 1999.
- [10] 田中和之, “画像修復の情報統計物理,” 特集 C 知識情報処理の統計力学のアプローチ, 数理科学, vol.e7, no.12, pp.21-27, December 1999.
- [11] R. Morina “On the hierarchical Bayesian approach to image restoration: application to astronomical images,” IEEE Trans. Image Processing, Vol. 8, No.2, pp.231–245, Feb.1999

- [12] R. Morina, K. Katsaggelos “Bayesian and regularization methods for hyperparameter estimation in image restoration,” IEEE Trans. Image Processing, Vol. 8, No.2, pp.231–245, Feb.1999
- [13] J. Inoue and K. Tanaka, “Dynamics of the maximum marginal likelihood hyper-parameter estimation in image restoration : Gradient descent versus expectation and maximization algorithm ,” Physical Review E, vol.65 pp. 016125-1 – 016125-11. 2002.
- [14] 井上純一, 田中和之 “ベイズ統計に基づく画像修復アルゴリズムの動的性質,” 信学技報, PRMU2000-124, pp.33–40, Dec.2000
- [15] H.Nishimori and K. Y. M .Wong, “Statistical mechanics of image restoration and error-correcting codes,” Physical Review E, vol.60, no.1, pp.132–144, July.1999
- [16] 西森秀稔 “有限温度の画像修復,” 物性研究, vol.73, pp.850–857, Feb.2000
- [17] 田中和之, 井上純一 “可解確率場モデルに基づく濃淡画像の画像修復のベイズ統計的アプローチ,” 信学技報, PRMU2000-125, pp.41–48, Dec.2000
- [18] K. Tanaka and J. Inoue, “Maximum likelihood hyperparameter estimation for solvable Markov random field model in image restoration,” IEICE Trans. Inf. & Syst, vol.E84-D, no.3, pp.546-557, 2002.
- [19] H.Derin, H.Elliott, R.Cristi and D.Geman, “Bayes smoothing algorithms for segmentation of binary images modeled by Markov random fields,” IEEE Trans. Pattern Anal. & Machine Intell.,vol.6, no.6, pp.707–720, Nov. 1984
- [20] S.Geman and D.Geman, “Stochastic relaxation, Gibbs distribution and the Bayesian restoration of images,” IEEE Trans. Pattern Anal. & Machine Intell., vol.6, no.6, pp.721–741, Nov. 1984
- [21] Jun-ichi Inoue, “Image restoration using the chiral Potts spin glass,” Physical Review E, vol.60, no.3, pp.2547–2553, Apr.1999

- [22] Jun-ichi Inoue, “Application of the quantum spin glass theory to image restoration,” *Physical Review E*, vol.63, pp.046114–046124, Mar.2001
- [23] Jun-ichi Inoue, “Image restoration using the Q-Ising spin glass,” *Physical Review E*, vol.64, pp.036121–036139, Aug.2001
- [24] J. Inoue and K. Tanaka, “Dynamical properties of image restoration and hyper-parameter estimation,” *Neural Networks for Signal Processing XI*, David J. Miller et al (Eds), pp. 383-392 (2001).
- [25] F. Rosenblatt, “Principles of neurodynamics - Perceptrons and the theory of brain mechanisms,” *Spartan*, (1961).
- [26] J.J. Hopfield, “Neurons with graded response have collective computational properties like those of two-state neurons,” *Proc. Natl. Acad. Sci. USA*, vol.81, May., pp.3088–3092, (1984).
- [27] P.M.C. de Oliveira, “Computer simulations of neural networks,” *Computer in Physics*, vol.11, no.5, Sep.–Oct., pp.443–449, (1997).
- [28] Y. Dai and Y. Nakano, “Recognition of facial images with low resolution using a Hopfield memory model,” *Pattern Recognition*, vol.31, no.2, Sep., pp.159–167, (1998).
- [29] J.J.Hopfield, D.I.Feinstein, and R.G.Palmer, “‘Unlearning’ has a stabilizing effect in collective memories,” *Nature*, vol.304, no.14, July., (1983).
- [30] M. Kimura and R. Nakano, “Learnig dynamical systems by recurrent neural networks from orbits,” *Neural Networks*, vol.11, pp.1989–1599, (1998).
- [31] Eric Hartman, “A high storage capacity neural network content-addressable memory,” *Network*, vol.2, pp.315–334, (1991).
- [32] J.Hertz, A.Krogh and R.G.Palmer, “Introduction to the theory of Neural computation,” Santa Fe Institute, Preseus Books Publishing, (1991).

- [33] S. Amari and K. Maginu :Statistical neurodynamics of associative memory; *Neural Networks*,Vol.1, pp.63–73 (1988)
- [34] D. Amit, H. Gutfreund, and H. Sompolinsky, :Statistical mechanics of neural networks near Saturation; *Annals of Physics*,173, pp.30–67 (1987)
- [35] J. Bruck and V.P. Roychowdhury, :On the number of spurious memories in the Hopfield model; *IEEE Trans. Information Theory*, Vol.36, No.2, pp.393–397, Mar (1990)
- [36] T. Fukai and M. Shiino :Comparative study of spurious-state distribution in analogue neural networks and Boltzmann machine; *Journal of Physics A:Mathematical and General*,Vol.25, pp.2873-2887 (1992)
- [37] J.J. Hopfield, :Neural networks and physical system with emergent collective computational abilities; *Proc. Natl. Acad. Sci. USA*, Vol.79, April., pp.2554–2558 (1982)
- [38] D.E. Rumelhart, G.E. Hinton, and R.J. Williams, :Learning representations by back-propagating errors; *Nature*,323, pp.533–536 (1986)
- [39] N. Surlas, :Spin-glass models as error-correcting codes; *Nature*,339, 693–695, (1989)
- [40] 西森 秀俊, “スピングラス理論と情報統計力学,” 岩波書店, (1999)
- [41] 利根川 進, :Use of knockout mice for the studies on neural plasticity and memory; 第 23 回日本神経科学大会, 招待講演, 横浜, 9 月 (2000)
- [42] 西森 秀稔, “ニューラルネットワークの統計力学,” 丸善株式会社, (1995).
- [43] 麻生 英樹, “ニューラルネットワーク情報処理,” 産業図書, (1988).
- [44] 上坂 吉則, “ニューロンコンピューティングの数学的基礎,” 近代科学社, (1993).