重畳エネルギー関数による多層パーセプトロンの冗長性削減*

高橋 隆史† 徳永 隆治‡

1997 年 3 月

概要

多層パーセプトロンの学習理論とその応用において，結合や素子の冗長性が，汎化能力の低下を引き起こし，内部表現を不明確にしてその解析を困難にすることが知られている．本論文では，出力計算に用いる中間層素子数の異なる誤差関数を複数個加算した簡明なエネルギー関数 (重畳エネルギー関数) を検討し，このエネルギー関数を用いた学習によって冗長性の低いパーセプトロンが構成され，中間層素子が寄与順に並んだ明確な内部表現が得られることを提示する．さらに，汎化能力改善における有効性について論議する．

1 まえがき

Rumelhart[1] による誤差逆伝搬学習則の提案以降，多層パーセプトロンの学習理論とその応用が幅広く研究されている．その理論的側面においては 3 層パーセプトロンの近似万能性が証明されており，十分な数の中間層素子を有するパーセプトロンは任意の関数を近似できることが知られている [2, 3, 4]．しかし，実利的側面においては，中間層素子の冗長性が未学習データに対する処理能力 (汎化能力) の低下を引き起こし，内部表現の解析を困難にするという問題点も指摘されている [5]．この視点において多層パーセプトロンの冗長性を削減する手法が検討されており，

1. 異なる素子数や結合数で構成されたパーセプトロンの中で最適なものを選択する方法 [6, 7]
2. 十分な素子および結合を有するパーセプトロンの構造を変化させ冗長性を削減する方法 [8]
   - (2-a) 誤差逆伝搬学習中あるいは学習後に，適当な基準に従って不要な素子と結合を削除する方式 [9, 10]
   - (2-b) 自由度に関する制約項を付加したエネルギー関数によって冗長な結合を自律的に減弱させる方式 [11, 12, 5]

等が提案されている．

他方，筆者ら [13] は，3 層線形パーセプトロンの恒等写像学習において，KL 変換に一致した内部表現を獲得するための簡明なエネルギー関数 (重畳エネルギー関数) を検討しており，中間層素子の結合を寄与の大きさに従って強化することで冗長性の低い内部表現が獲得できることを指摘している．特に，重畳エネルギー関数にはパーセプトロンの構造上の制約はないため，上述のカテゴリ (2-b) の一方式として一般の学習課題に適用できる．そこで本研究では，一変数の関数近似問題および二次元パターン分類問題を例に，重畳エネルギー関数に基づく学習則によって，冗長性の低い内部表現を持つパーセプトロンが構成できることを報告する．

以降，2において重畳エネルギー関数に基づく学習則について説明し，3においてその有効性を数値実験により示す．さらに，4において冗長性の削減と汎化能力との関連について議論し，5において結論と今後の課題を述べる．

2 重畳エネルギー関数に基づく学習則

2.1 諸定義

重畳エネルギー関数に基づく学習則の導入に際していくつかの定義を行う．

定義 1 部分パーセプトロン

入力層素子 N 個，中間層素子 H 個，出力層素子 M 個から成る *3* 層パーセプトロンにおいて[1]，入力

$$\boldsymbol{x} = {}^T[x_1, x_2, \ldots, x_N] \tag{1}$$

が与えられたときの中間層素子の出力を

$$\boldsymbol{u} = {}^T[u_1, u_2, \ldots, u_H] \tag{2}$$

とする．$u_h (h = 1, 2, \ldots, H)$ は

$$u_h = \sigma\left(\sum_{n=1}^{N} v_{h,n} x_n + v_{h,0}\right) \tag{3}$$

で与えられる．ただし，$v_{h,n}$ は中間層第 h 素子と入力層第 n 素子の結合，$v_{h,0}$ はしきい値を表し，$\sigma(\cdot)$ は中

---

*電子情報通信学会和文論文誌 D-II に掲載決定

†筑波大学大学院博士課程工学研究科 ume@chaos.is.tsukuba.ac.jp

‡筑波大学電子・情報工学系

[1] 以降の議論はすべて 3 層パーセプトロンを仮定するが，より多層のパーセプトロンに関しても容易に拡張可能である．

間層素子の出力関数である．一般のパーセプトロンは u_1 から u_H まで全ての出力を考慮するが， u_1 から u_i までを考慮し，残りの素子の出力 $u_{i+1}, \ldots, u_H$ を 0 で置き換えて得られるパーセプトロンを第 i 部分パーセプトロンと呼ぶ (図 1参照)．

第 i 部分パーセプトロンの出力は，

$$\boldsymbol{y}^i = {}^T[y_1^i, y_2^i, \ldots, y_M^i] \quad (i = 1, 2, \ldots, H) \tag{4}$$

ただし， $y_m^i (m = 1, 2, \ldots, M)$ は

$$y_m^i = \rho\left(\sum_{h=1}^{i} w_{m,h} u_h + w_{m,0}\right) \tag{5}$$

で与えられる．ここで， $w_{m,h}$ は出力層第 m 素子と中間層第 h 素子の結合， $w_{m,0}$ はしきい値を表し， $\rho(\cdot)$ は出力層素子の出力関数である．特に，第 H 部分パーセプトロンの出力 $\boldsymbol{y}^H$ は全ての中間層素子を用いる通常のパーセプトロンの出力に一致することに注意する．

定義 2 部分エネルギー関数

T 個の入出力データ対

$$\{\tilde{\boldsymbol{x}}_t \in \boldsymbol{R}^N, \tilde{\boldsymbol{y}}_t \in \boldsymbol{R}^M\}_{t=1,2,\ldots,T} \tag{6}$$

で教師データを与える．入力 $\tilde{\boldsymbol{x}}_t$ に対する第 i 部分パーセプトロンの出力を $\boldsymbol{y}_t^i$ とするとき，誤差関数

$$E_i = \frac{1}{2T}\sum_{t=1}^{T} ||\tilde{\boldsymbol{y}}_t - \boldsymbol{y}_t^i||^2 \quad (i = 1, 2, \ldots, H) \tag{7}$$

を第 i 部分エネルギー関数と呼ぶ．この値は (第 H 部分) パーセプトロンが目標の出力値を生成する上で第 i 部分パーセプトロンがどれだけ貢献しているかを示す指標と考えられる．

定義 3 重畳エネルギー関数

部分エネルギー関数 E_i の重み付き和

$$F = \sum_{i=1}^{H} \beta_i E_i \tag{8}$$

を重畳されたエネルギー関数 (重畳エネルギー関数) と呼ぶ． β_i は正の定数である (β_i の設定に関しては付録で論議する)．

2.2 重畳エネルギー関数による学習則

一般の誤差逆伝搬学習の場合，すべての素子を用いて得られる誤差 E_H のみが減少するように学習が進行する．しかし，重畳エネルギー関数を用いた学習では，第 i 中間層素子に関与する結合の修正量は $E_i, \ldots, E_H$ より計算されるが，第 $i-1$ 素子の結合については，これらに加えて i 番目以下の素子を用いないで得られる誤差 E_{i-1} も減少させるように修正される．即ち，番号の小さな素子に関する結合ほど，少ない素子数で目標出力を生成するように学習が進行する．以上より，上位の素子ほど全体の誤差 E_H の減少に対して大きく寄与し，素子の番号順に重要な特徴が抽出されると考えられる．

式 (8) で与えられる重畳エネルギー関数に基づく学習則を導出する．例えば，出力層素子の結合 $w_{m,h}$ の修正量 $\Delta w_{m,h}$ を考慮すると， F に関する最急降下法より直ちに

$$\Delta w_{m,h} = -\eta\frac{\partial F}{\partial w_{m,h}} = -\eta\sum_{i=1}^{H}\beta_i\frac{\partial E_i}{\partial w_{m,h}} \tag{9}$$

が導かれる．ただし η は学習定数である． $i < h$ のとき

$$\frac{\partial E_i}{\partial w_{m,h}} = 0 \tag{10}$$

であるから，式 (9) は

$$\Delta w_{m,h} = -\eta\sum_{i=h}^{H}\beta_i\frac{\partial E_i}{\partial w_{m,h}} \tag{11}$$

となる．従って， $\Delta w_{m,h}$ は対応する各部分パーセプトロンに通常の誤差逆伝搬学習則を適用したときの結合修正量の和で与えられる．入力層側の結合 $v_{h,n}$ およびしきい値 $v_{h,0}$ についても同様である．一方，しきい値 $w_{m,0}$ はすべての部分パーセプトロンにおいて考慮されるため，その学習則は

$$\Delta w_{m,0} = -\eta\sum_{i=1}^{H}\beta_i\frac{\partial E_i}{\partial w_{m,0}} \tag{12}$$

で与えられる．

3 数値実験

重畳エネルギー関数の効果を示すためにいくつかの数値実験を行う．

3.1 実験 1 一変数関数近似問題

学習対象として連続関数 [14]

$$f : x \mapsto \frac{(x-2)(2x+1)}{x^2+1} \tag{13}$$

を考慮する．入力データ $\tilde{x}_t$ は区間 $[-8, 8]$ において刻み幅 0.3 で等間隔に選択し，教師データ $\tilde{y}_t$ は $f(\tilde{x}_t)$ に平均 0，分散 0.1 のガウシアンノイズを加えたものとする (図 2参照)．実験には

$$(N, H, M) = (1, 12, 1) \tag{14}$$

の 3 層パーセプトロンを用いる．また，中間層の出力関数はシグモイド関数

$$\sigma(x) = \tanh(x) \tag{15}$$

とし，$w_{m,0} \equiv 0 (m=1)$ のもとで出力層の出力関数は線形

$$\rho(x) = x \tag{16}$$

とする．学習に用いる諸パラメータを表 1に示す．

図 3(a) は，重畳エネルギー関数を用いて構成された結合の初期値の異なる 10 個のパーセプトロンの部分エネルギー関数 E_i を示している．初期値に依存することなく，部分エネルギー関数は中間層素子数に対してほぼ単調減少し，7 番目の中間層素子で収束している．即ち，8 から 12 番目の中間層素子は関数 (13) の近似問題において冗長であると考えられる．次に，この結果が重畳エネルギー関数に固有の効果によることを示すため，通常のエネルギー関数に基づく結果と比較する．前者同様に 10 個の異なるパーセプトロンを構成し，学習後に $i = H, H-1, \ldots, 2$ として操作

> 「1 番目から i 番目の中間層素子のうち，一つ削除したときに E_i に対する誤差の増加が最も少ないものを選択して，その素子を新たに i 番目とする．」

を繰り返して貢献度の順に素子の置換を行う．図 3(b) は，この操作によって得られる部分エネルギー関数 E_i を示している．中間層素子数の増加に伴い E_i が減少する傾向は見られるが，$i = 1, 2, \ldots, H-1$ の部分パーセプトロンでは誤差の収束は確認できない．即ち内部表現は冗長な分散表現となっていると考えられる．

両者の内部表現の違いを検証するため，部分パーセプトロンの出力 y^i を図 4に示す．重畳エネルギー関数によるパーセプトロンでは少数の中間層素子で概形が表現され，素子の追加に従って徐々に細部が近似されるが，通常のエネルギー関数によるパーセプトロンでは少数の中間層素子での近似精度は悪く，それぞれの役割が不明瞭である．

3.2　実験 2　二次元パターン分類問題

次に，二次元のらせん上の点にノイズを加えて生成される教師データ (図 5参照) を用いた二つのクラスのパターン分類問題を考慮する．パーセプトロンの構成は

$$(N, H, M) = (2, 10, 2) \tag{17}$$

とし，中間層および出力層の出力関数はシグモイド関数

$$\sigma(x) = \rho(x) = \frac{1}{1 + \exp(-x)} \tag{18}$$

とした．また，出力の教師データは $\{0, 1\}$ の代わりに $\{0.1, 0.9\}$ を選択し，表 2に示された諸パラメータを利用した．重畳エネルギー関数を用いて結合初期値の異なる 10 個のパーセプトロンを構成し，それぞれの部分パーセプトロンについてテストデータに対する認識率を求める．テストデータは各点の座標値に加えるノイズの値を変更して生成し，認識率は，二つの出力素子のうち大きな値の素子の示すクラスが正しいクラスに一致した割合として算出する (図 6(a) 参照)．認識率は 7 番目の中間層素子でほぼ収束しており，8 から 10 番目までの中間層素子は冗長なものと考えられる．

結果の比較のため，通常のエネルギー関数を用いて同様に 10 個のパーセプトロンを構成する (表 2参照)．ただし，50,000 回の反復計算では学習が停滞して十分な性能のパーセプトロンが構成できなかったため，反復回数を 100,000 回とした．実験 1 と同様に中間層素子の置換を行って，部分パーセプトロンの認識率を計算する (図 6(b) 参照)．3 個のパーセプトロンでは 100,000 回の反復でも十分な認識率は得られなかった．それ以外についても $i = 1, 2, \ldots, H-1$ での認識率の収束は見られず，内部表現はやはり冗長であると考えられる．

図 8は，重畳エネルギー関数および通常のエネルギー関数で構成されたパーセプトロンについて，部分パーセプトロンを用いて入力平面の領域分割を行った結果を示す．白丸の教師データをクラス 1，黒丸をクラス 2 として対応する部分パーセプトロン出力を y_1^i および y_2^i と定めるとき，図中の実線は

$$y_1^i = y_2^i \tag{19}$$

を満たす二つのクラスの境界を示す[2]．また，条件

白色　$y_1^i > y_2^i$ かつ $y_1^i \geq 0.75$
濃色　$y_1^i < y_2^i$ かつ $y_2^i \geq 0.75$
淡色　上記以外

に従って領域を分割し，塗分けを行った．さらに，図中の破線は各中間層素子の識別線

$$\begin{aligned} l_h &= \{(x_1, x_2) | u_h = 0.5\} \\ &= \left\{ (x_1, x_2) \middle| \sigma\left(\sum_{n=1}^{2} v_{h,n} x_n + v_{h,0}\right) = 0.5 \right\} \\ &= \{(x_1, x_2) | v_{h,1} x_1 + v_{h,2} x_2 + v_{h,0} = 0\} \end{aligned} \tag{20}$$

を示している．重畳エネルギー関数の場合，上位 4 つの中間層素子で大まかな分類を行い (図 8(a))，さらに 2 素子の追加で主に中心領域の分類が修正されて概形が完成し (同 (b))，残りの素子は細部の修正を行っている (同 (c))．一方通常のエネルギー関数の場合，上位の中間層素子のみで構成された部分パーセプトロンでは概形は得られず (図 8(d),(e))，各中間層素子の役割が明確にならない．また，l_5 と l_7 の識別線 (同 (e),(f)) がほぼ平行であり，明らかに冗長な素子が存在することが理解される．

[2]厳密には，式 (19) を満たす入力 (x_1, x_2) の集合は領域となる．しかし本実験の場合，図に示すスケールではそのような領域を確認できないため，境界領域を曲線として扱う．

4　議論

重畳エネルギー関数に基づく学習では，冗長性の低い内部表現が自己組織され，汎化能力が良好に保持される可能性がある．これを調べるため，実験 1 で構成したパーセプトロンを用い，真の関数値 (13) に対する誤差を指標として汎化能力を評価する．なお，テストデータは区間 $[-10, 10]$ においてきざみ幅 0.1 で等間隔に選択した．

重畳エネルギー関数および通常のエネルギー関数により構成されたパーセプトロンの誤差を図 7(a) および (b) にそれぞれ示す．重畳エネルギー関数の場合，図 3(a) と同様にテストデータに対する誤差も 7 個程度の中間層素子数で収束している．特に，10 個の第 12 部分パーセプトロンの平均誤差 (図中破線) は 0.0107 となっており，通常のエネルギー関数の平均 0.0221 の半分程度となっている．従って，重畳エネルギー関数を用いると，内部表現の冗長性削減と共に汎化能力が向上すると考えられる．

この結果を他の汎化能力改善手法と比較するため，正則化項 (忘却項) を付加したエネルギー関数 [15, 16]

$$G = E_H + \frac{\gamma}{2}\left(\sum_{h=1}^{H}\sum_{n=0}^{N} v_{h,n}^2 + \sum_{m=1}^{M}\sum_{h=0}^{H} w_{m,h}^2\right) \quad (21)$$

を用いて追実験を行う．図 7(c) は．実験 1 と同様の条件で 10 個のパーセプトロンを構成し，中間層素子の置換後にテストデータに対して求められた誤差を示している[3]．第 12 部分パーセプトロンの平均誤差では重畳エネルギー関数と同程度の性能が得られている．しかし，図 7(b) と同 (c) を比較すると，正則化項を付加した場合，通常のエネルギー関数よりも中間層素子の利用

に偏りが少なく，分散表現が強まっていることが理解できる．従って，重畳エネルギー関数による学習と通常のエネルギー関数に正則化項を併用する学習では，全く異なる内部表現によって汎化能力が向上されている．

5　むすび

重畳エネルギー関数に基づく学習によって，冗長性の低いパーセプトロンが自己組織されることを示した．得られる内部表現は少数の中間層素子に集中し，寄与順に並んだ明確なものとなる．そのため，各素子についての貢献度の評価が可能であり，パーセプトロンの解析が容易となる．また，冗長性削減の効果によって汎化能力の優れたパーセプトロンが構成される可能性を指摘した．

重畳エネルギー関数を用いた冗長性削減方式は，エネルギー関数の変更によるものであり，従来提案されている冗長性削減手法との併用が可能である．この観点から，より性能の高いパーセプトロンの構築法を検討することが一つの課題となる．また，β_i の選択が学習性能や内部表現に与える影響についても今後検討の必要があると考えられる．他方，汎化能力の向上については，統計的手法 [6, 16, 7] との比較を含めてより詳細な考察が必要であり，従って重畳エネルギー関数の学習性能に関した理論的検討も今後の課題とする．

[3]係数 γ については，様々な値を用いて試行錯誤を行い，テストデータに対する誤差が最小となる値 $\gamma = 0.32$ を選択した

参考文献

[1] D. E. Rumelhart, G. E. Hinton and R. J. Williams, “Learning internal representations by error propagation,” in Parallel Distributed Processing vol.1, eds. J. L. McCleland, D. E. Rumelhart, and The PDP Research group, chap.8, MIT Press, Cambridge, 1986.

[2] K. Funahashi, “On the approximate realization of continuous mappings by neural networks,” Neural Networks, vol.2, pp.183–192, 1989.

[3] K. Hornik, M. Stinchcombe and H. White, “Multilayer feedforward networks are universal approximators,” Neural Networks, vol.2, pp.359–366, 1989.

[4] G. Cybenko, “Approximation by superpositions of a sigmoidal function,” Mathematics of Control, Signals, and Systems, vol.2, no.4, pp.303–314, 1989.

[5] M. Ishikawa, “Structural learning with forgetting,” Neural Networks, vol.9, no.3, pp.509–521, 1996.

[6] 栗田多喜夫, “情報量基準による 3 層ニューラルネットの隠れ層のユニット数の決定法,” 信学論 (D-II), vol.J73-D-II, no.8, pp.1872–1878, 1990.

[7] N. Murata, S. Yoshizawa and S. Amari, “Network information criterion — Determining the number of hidden units for an artificial neural network model,” IEEE Trans. on NN, vol.5, no.6, pp.865–872, 1994.

[8] R. Reed, “Pruning algorithms — A survey,” IEEE Trans. on NN, vol.4, no.5, pp.740–747, 1993.

[9] Y. Le Cun, J. S. Denker and S. A. Solla, “Optimal brain damage,” in Advances in neural information processing systems 2, pp.598–605, 1990.

[10] 萩原将文, “淘汰機能を有するバックプロパゲーション — 学習回数の低減と中間層ユニットの削減法 —,” 信学論 (D-II), vol.J74-D-II, no.6, pp.812–818, 1991.

[11] S. J. Hanson and L. Y. Pratt, "Comparing biases for minimal network construction with back-propagation," in Advances in neural information processing systems 1, pp.177–185, 1989.

[12] Y. Chauvin, "A back-propagation algorithm with optimal use of hidden units," in Advances in neural information processing systems 1, pp.519–526, 1989.

[13] 高橋隆史, 徳永隆治, 平井有三, "KL 変換を実現する 3 層線形パーセプトロンの教師付き学習則 / Baldi-Hornik の定理の拡張," 信学論 (D-II), 掲載決定.

[14] M. H. Hassoun, "Fundamentals of artificial neural networks," MIT Press, pp.221–226, 1995.

[15] A. Krogh and J. A. Hertz, "A simple weight decay can improve generalization," in Advances in neural information processing systems 4, pp.951–957, 1992.

[16] D. J. C. MacKay, "Bayesian methods for adaptive models," Ph. D. Thesis, California Institute of Technology, 1992.

係数 β_i の選択

線形パーセプトロンの恒等写像学習に重畳エネルギー関数を適用する場合，理論上は正定数 β_i を任意に選択できる [13]．しかし，実際に線形および非線形の多層パーセプトロンに適用する際には学習効率が係数 β_i に依存して変化する．特に，β_i は $\sum_{i=1}^{H} \beta_i = 1$ としても一般性を失わないため，以降ではこれを仮定して議論する．

$w_{m,h}$ を例に，すべての部分エネルギー関数に関するこの結合の修正量がほぼ等しい場合を考慮する．すなわち，

$$\frac{\partial E_h}{\partial w_{m,h}} \approx \frac{\partial E_{h+1}}{\partial w_{m,h}} \approx \cdots \approx \frac{\partial E_H}{\partial w_{m,h}} \tag{22}$$

を仮定する (ただし $h = 1, 2, \ldots, H$ である．)．このとき，式 (11) より

$$\Delta w_{m,h} \approx -\eta \left(\sum_{i=h}^{H} \beta_i \right) \frac{\partial E_H}{\partial w_{m,h}} \tag{23}$$

であるから，$w_{m,h}$ の学習定数は実質的に $w_{m,1}$ の $\sum_{i=h}^{H} \beta_i$ ($\leq \sum_{i=1}^{H} \beta_i = 1$) 倍となる．他の結合についても同様の議論が成り立つ．ゆえに，例えば β_i を i によらず一定とすると，下位の中間層素子の結合修正量が上位のものに比べて非常に小さくなる．その結果，各部分エネルギー関数の間の差が大きくならず，各素子の寄与が偏りにくくなる．これを回避するためには，$\{\beta_i\}_{i=1,2,\ldots,H}$ に単調増加性を仮定し，実質的な学習定数の差を軽減することが必要となる．

特に，本研究では経験的に β_i を

$$\beta_i = \frac{i^2}{\sum_{j=1}^{H} j^2} = \frac{6i^2}{H(H+1)(2H+1)} \tag{24}$$

と設定した．

表 1: 関数近似問題の実験条件

initial weight	uniform random number over $[-0.1, 0.1]$
learning constant	0.01
momentum	0
iteration	50,000 epoch

表 2: パターン分類問題の実験条件

	superposed	conventional
initial weight	uniform random number over $[-0.1, 0.1]$	
learning constant	0.5	
momentum	0.9	
iteration	50,000 epoch	100,000 epoch

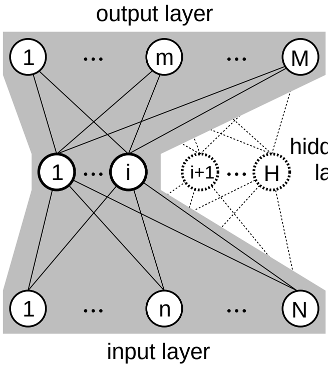

図 1: パーセプトロンおよび部分パーセプトロン：実線は第 i 部分パーセプトロンを示し，破線を含めてパーセプトロン全体を示している．

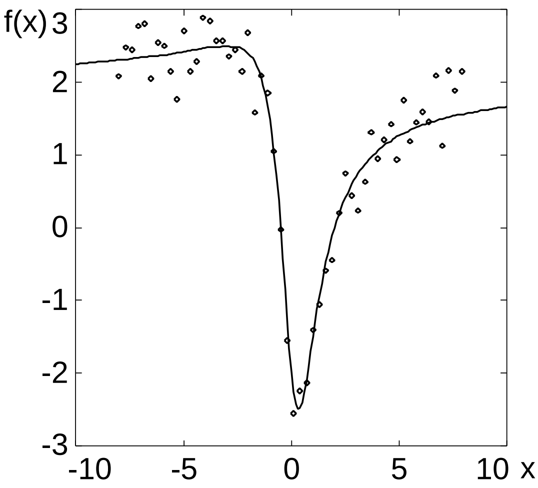

図 2: 関数 (13) と教師データ (実線: 関数 (13)，点: ノイズの付加された教師データ)

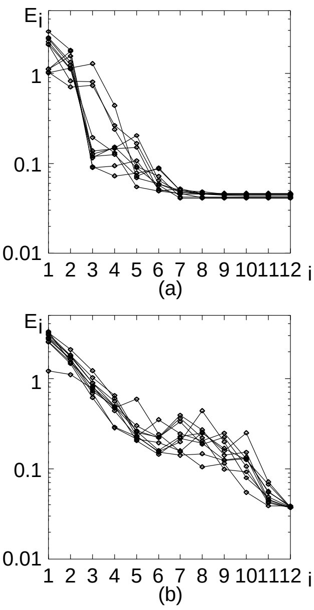

図 3: 関数近似問題における部分エネルギー関数 ((a) 重畳エネルギー関数，(b) 通常のエネルギー関数，横軸: 中間層素子数 i，縦軸: 第 i 部分エネルギー関数 E_i)

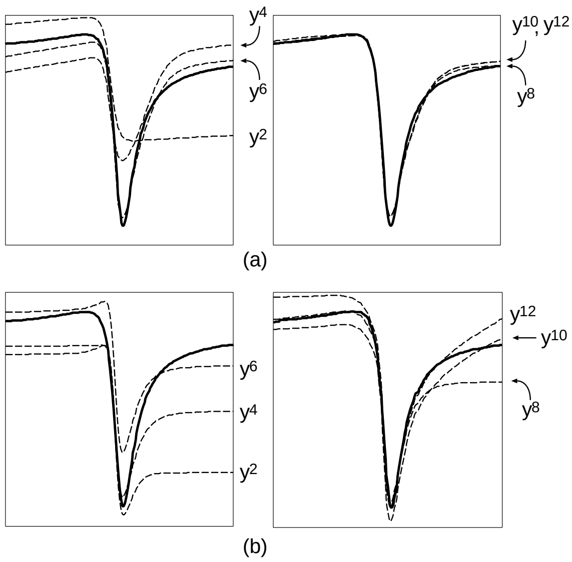

図 4: 関数近似問題における部分パーセプトロンの出力 ((a) 重畳エネルギー関数，(b) 通常のエネルギー関数，実線: 真の関数曲線，破線: 第 i 部分パーセプトロンの出力 $y^i(i=2,4,6,8,10,12)$)

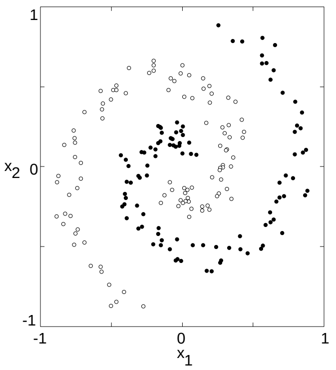

図 5: パターン分類問題の教師データ (白丸および黒丸：二つのクラス，座標 (x_1, x_2)：入力値)

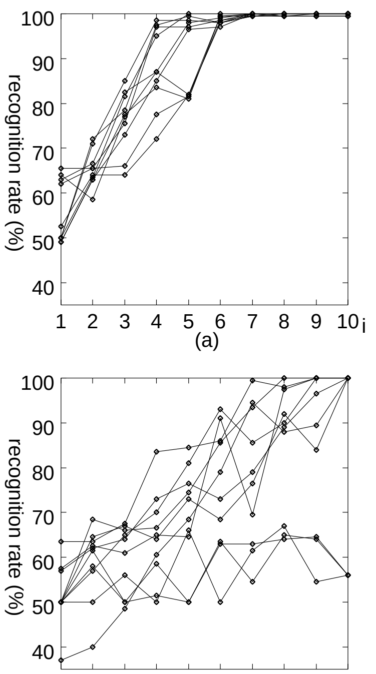

図 6: パターン分類問題における部分パーセプトロンの認識率 ((a) 重畳エネルギー関数，(b) 通常のエネルギー関数，横軸: 中間層素子数 i，縦軸: 第 i 部分パーセプトロンの認識率)

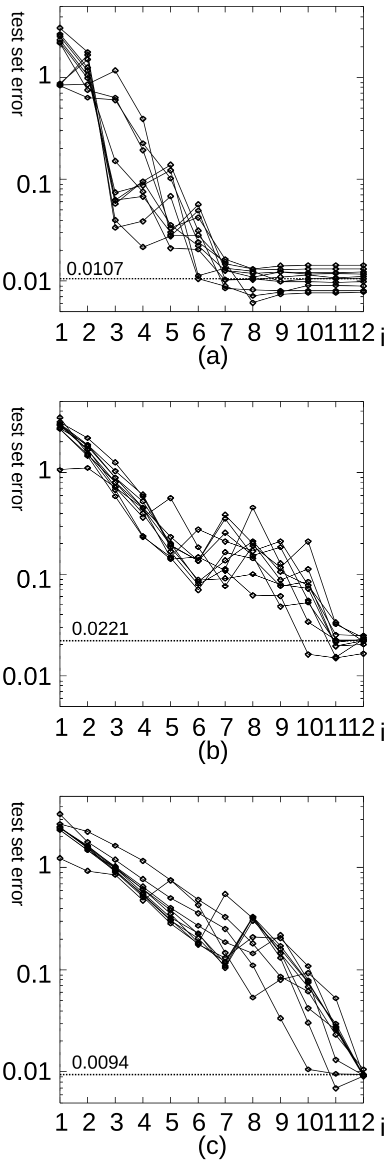

図 7: テストデータに対する誤差 ((a) 重畳エネルギー関数，(b) 通常のエネルギー関数，(c) 正則化項を付与した通常のエネルギー関数，横軸: 中間層素子数 i，縦軸: 第 i 部分パーセプトロンの誤差，点線：第 12 部分パーセプトロンの平均値．

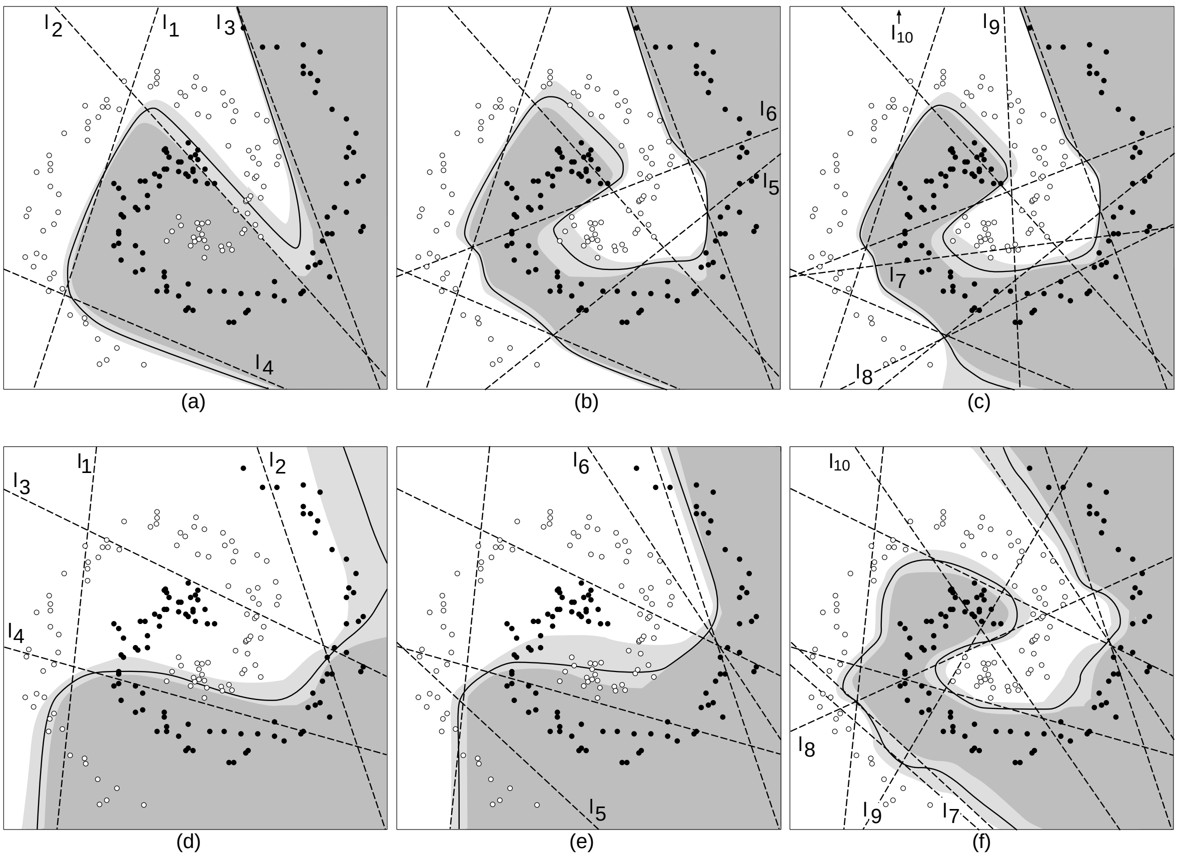

図 8: 部分パーセプトロンによる入力平面の領域分割 ((a)–(c): 重畳エネルギー関数，(d)–(f): 通常のエネルギー関数，(a),(d) $i = 4$，(b),(e) $i = 6$，(c),(f) $i = 10$ の部分パーセプトロンに対応．実線：分類境界，破線：識別線 l_h．)