

UNIVERSIDADE ESTADUAL DE CAMPINAS
FACULDADE DE ENGENHARIA ELÉTRICA
DEPARTAMENTO DE ENGENHARIA DE COMPUTAÇÃO E AUTOMAÇÃO INDUSTRIAL

DETERMINAÇÃO DOS PARÂMETROS DE ORDEM DE REDES NEURAIS PELO MÉTODO DOS CUMULANTES

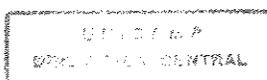
Este exemplar corresponde à redação final da tese defendida por _____
_____ pela Comissão Julgadora em _____
Orientador _____

Orientador:
Prof. Dr. Márcio Luiz de Andrade Netto

Aluno:
José Roberto Campanha

Tese apresentada à Faculdade de Engenharia Elétrica da Universidade Estadual de Campinas, UNICAMP, como parte dos requisitos exigidos para a obtenção do título de Doutor em Engenharia Elétrica.

JULHO 94



Para

Seu Guilherme , a Dona Lila e a Kassima

Agradecimentos

Gostaria de agradecer ao meu orientador Prof. Dr. Márcio Luiz de Andrade Netto pela rara virtude de acolher e saber incentivar pessoas e idéias advindas de outras áreas. Ao Prof. Dr. Aldo Tancredo pela discussão quase que diária de alguns dos temas abordados nesse trabalho e ao Prof. Dr. José Fernando Fontanari que com seu profundo conhecimento de Mecânica Estatística de Redes Neurais esclarecia, sempre gentilmente minhas inúmeras dúvidas.

Gostaria de agradecer também ao Departamento de Física da UNESP de Rio Claro por fornecer todas as condições para a realização desse trabalho.

E ainda ao Gustavo H. R. dos Santos que ajudou-me nos problemas computacionais e na impressão desse trabalho.

Muito obrigado

ÍNDICE

1 - INTRODUÇÃO.....	01
2 - REDES NEURAIS.....	04
2.1 - Fatos básicos sobre neurônios e redes neurais.....	04
2.2 - Principais modelos de redes neurais (1943 -1982).....	12
3 - MODELO DE HOPFIELD.....	32
3.1 - O trabalho de Hopfield.....	32
3.2 - Padrões aprendidos como atratores da rede	37
3.3 - Analogia spin-neurônio	41
4 - MÉTODO DOS CUMULANTES.....	48
4.1 - Método dos cumulantes.....	48
4.2 - Aplicação do método ao modelo de Hopfield.....	56
4.3 - Diagrama de fase da rede de Hopfield.....	61
4.4 - Aplicações a outros modelos.....	65
A - Modelo de rede neural com esquecimento.....	65
B - Redes neurais com baixo nível de atividade.....	67
5 - MODELO PROPOSTO.....	70
5.1 - Cálculo dos parâmetros de ordem do modelo	70
5.2 - Resultados do modelo	73

6 - CONCLUSÕES	79
APÊNDICE A:	
Vidros de Spins.....	82
APÊNDICE B:	
Método das Réplicas.....	85
APÊNDICE C:	
Aplicação dos Métodos da Mecânica Estatística ao Modelo de Hopfield	87
APÊNDICE D:	
Aplicação dos Métodos da Mecânica Estatística ao Modelo Proposto	105
7 - BIBLIOGRAFIA.....	110

RESUMO

Propomos neste trabalho um método mais simples, para o cálculo dos parâmetros de ordem de redes neurais, baseados nos cumulantes de uma distribuição.

Aplicamos este método ao modelo de Hopfield, obtivemos seus parâmetros de ordem m, q e r e baseados nestes parâmetros, construímos seu diagrama de fase.

A seguir, estudamos com nosso método, um modelo de rede neural com esquecimento, e outra rede com baixo nível de atividade neural obtendo novamente, os mesmos resultados que os obtidos pelos métodos da Mecânica Estatística.

Propomos ainda um modelo de uma rede neural, no qual as interações entre os neurônios são dadas por:

$$J_{ij} = \frac{1}{N} \sum_{\mu=1}^N \xi_i^{\mu} \xi_j^{\mu} \cos[k(x_i - x_j)]$$

Trata-se de uma generalização da regra de Hebb, em que consideramos a variação espacial da interação.

A partir do método dos cumulantes calculamos os parâmetros de ordem do modelo e mostramos que para $k > \frac{\pi}{2L}$ não há recuperação dos padrões aprendidos pela rede.

A principal vantagem do método proposto é ser mais simples do que os métodos da Mecânica Estatística, e por isto acredito, ser acessível a profissionais de outras áreas.

*dois loucos no bairro
um passa os dias
chutando postes para ver se acendem*

*o outro passa as noites
apagando palavras
contra um papel branco*

*todo bairro tem um louco
que o bairro trata bem
só falta mais um pouco
pra eu ser tratado também*

Paulo Leminski

CAPÍTULO 1

INTRODUÇÃO

A compreensão do funcionamento do cérebro humano é um dos principais objetivos da ciência. O cérebro é a máquina mais poderosa conhecida do homem. O ato de pensar, lembrar, resolver problemas, inspirou muitos cientistas a tentarem imitar essas operações. Para isto utilizam-se computadores e programas que de forma muito simples procuram simular algumas das complexas operações do cérebro. Os resultados dessa pesquisa constituem o que denominamos de computação neural ou estudo das redes neurais.

Várias áreas da ciência têm contribuído no estudo das redes neurais. Uma destas áreas, a Mecânica Estatística, procura propriedades gerais das redes efetuando médias sobre o comportamento de um número muito grande de unidades de processamento (denominadas de neurônios) interligadas, as quais são supostas possuírem uma estrutura interna bastante simples.

Os resultados obtidos, tais como: a capacidade de armazenamento de padrões das redes, o efeito do ruído no armazenamento destes padrões, seus parâmetros de ordem e, principalmente, seu diagrama de fase, embora muito importantes ficam distantes de profissionais de outras áreas devido à complexidade inerente aos métodos para a sua obtenção.

O objetivo deste trabalho é aplicar técnicas mais simples da Estatística, o método dos cumulantes, para obter os parâmetros de ordem de uma rede neural e, a

partir daí, a confecção do diagrama de fase da rede, sem a utilização dos métodos da Mecânica Estatística. Espera-se assim, tornar acessível aos profissionais de outras áreas a confecção do diagrama de fase de redes neurais.

A origem do trabalho baseia-se no artigo " Thermodynamic properties of spin glasses " (1), cujos autores, M. Fibich, I. Riess e A. Aron reproduzem os resultados do modelo de Sherrington-KirkPatrick (modelo SK) de um vidro de spin através do método dos cumulantes. Procurou-se então, estender estes resultados para as redes neurais.

No segundo capítulo, introduziremos as noções básicas sobre neurônios e redes neurais. Trataremos da condução do estímulo nervoso, dos tipos de aprendizado e recuperação de padrões e da natureza distribuída e associativa das redes. Ainda neste capítulo abordaremos os principais modelos de redes neurais, desde o trabalho pioneiro de McCulloch e Pitts em 1943 até o modelo de Hopfield em 1982, considerado um dos trabalhos mais importantes da área por esclarecer parte das dúvidas em relação ao processo dinâmico das redes neurais.

Observaremos, por esta retrospectiva, que muitas das idéias de Hopfield estavam presentes em trabalhos de seus antecessores.

O trabalho de Hopfield será tratado no terceiro capítulo, sua dinâmica descrita, e será demonstrado que a função "energia" da rede decresce, ou permanece constante durante cada iteração do sistema, até a rede atingir uma configuração de energia mínima, que representa um padrão ensinado, os quais são os atratores da dinâmica da rede. Destacaremos ainda neste capítulo, o comportamento estocástico dos neurônios da rede e também a noção de temperatura, como um parâmetro que controla o nível de ruído da rede.

No quarto capítulo, introduziremos o método dos cumulantes e como podemos calcular o valor médio de uma função através de seus cumulantes. A partir daí, aplicaremos estes resultados ao modelo de Hopfield, obtendo seus parâmetros de ordem e, por fim, a construção de seu diagrama de fase.

Aplicaremos ainda o método a dois outros modelos, o de uma rede neural com esquecimento e a uma rede com baixo nível de atividade, a fim de mostrarmos a equivalência com os métodos da Mecânica Estatística na obtenção dos parâmetros de ordem.

No quinto capítulo, proporemos um modelo no qual as interações entre os neurônios possuem uma variação espacial dada por:

$$J_{ij} = \frac{1}{N} \sum_{\mu=1}^p \xi_i^{\mu} \xi_j^{\mu} \cos k(x_i - x_j) \quad (1.1)$$

Trata-se de um modelo mais geral que o de Hopfield pois reduz-se a este quando $k=0$, e foi resolvido neste trabalho para o caso $\alpha=0$ (o número de padrões ensinados permanece finito quando tomamos o limite termodinâmico) através do método dos cumulantes e da Mecânica Estatística (APÊNDICE D), obtendo-se os mesmos resultados. Ainda neste capítulo exporemos alguns resultados do modelo.

Finalmente, no sexto capítulo, abordaremos novos problemas aos quais o método dos cumulantes poderá ser aplicado e teceremos as considerações finais.

Foram ainda colocados como apêndices para melhor orientação dos leitores, uma breve introdução aos vidros de spins (APÊNDICE A), o método das réplicas (APÊNDICE B), o cálculo dos parâmetros de ordem para a rede de Hopfield através da Mecânica Estatística (APÊNDICE C), e o cálculo dos parâmetros de ordem do modelo proposto também via Mecânica Estatística (APÊNDICE D).

Evidentemente, os métodos da Mecânica Estatística, principalmente o método das réplicas são muito mais ricos conceitualmente que o método dos cumulantes. O método das réplicas foi o embrião de todo um desenvolvimento no estudo dos sistemas denominados desordenados que incluem, vidros de spins, redes neurais, polímeros, problemas de otimização etc....O método dos cumulantes é bem mais simples e esta é, acredito, sua principal virtude.

Neste trabalho procura-se, demonstrar a grande maioria dos resultados, esperando que tal fato amenize o caminho de outros que iniciem seus trabalhos nessa área tão maravilhosa quanto desafiante, que é o estudo das redes neurais.

CAPÍTULO 2

REDES NEURAIS

2.1 Fatos básicos sobre neurônios e redes neurais (2).

Atualmente existem numerosas observações experimentais sobre a estrutura e funcionamento do cérebro, e algumas idéias e modelos tratando com aspectos limitados das funções cerebrais. Existem ainda, muitas questões fundamentais não respondidas.

Nos últimos 50 anos, muitas tentativas foram feitas para simular o comportamento de algumas funções cerebrais e, responder parcialmente a algumas destas questões fundamentais.

As células nervosas (ou neurônios) foram primeiramente identificadas como constituintes celulares do cérebro por Cajal (3). Neurônios são células especializadas para intercomunicarem-se de forma muito rápida, principalmente através de correntes iônicas. Sua principal propriedade é a de gerar e transmitir impulsos elétricos. Estes impulsos foram descobertos cerca de oitenta anos atrás por Adrian (4) que denominou-os de "potencial de ação" ou respostas "tudo ou nada" uma vez que suas amplitudes são constantes. Em geral os neurônios emitem tais pulsos a uma taxa proporcional ao estímulo. Foram Cajal e Golgi (5) quem primeiro descobriram a imensa complexidade dos circuitos neurais. Estudos da neuroanatomia indicam um grau de

interconectividade, de várias ordens de magnitude maior que aquele encontrado mesmo nos mais avançados supercomputadores.

Os neurônios possuem, além de um corpo celular revestido por uma membrana, uma complexa rede de prolongamentos denominados dendritos, (FIGURA 1), que são ramificados e especializados em receber sinais de outros neurônios através das sinapses (6) e ainda, o axônio que é em geral bastante longo e que serve para conduzir o estímulo nervoso para regiões mais distantes.

Por exemplo, algumas células, as células de Purkinge do cerebellum, possuem cerca de 200.000 sinapses (7).

As redes assim resultantes são muito complexas. Existe uma estimativa de que o cérebro humano possui cerca de 10^{10} neurônios (8). Existem muitos tipos diferentes de neurônios. Szentágothai, enumerou cerca de 250 tipos morfológicos diferentes de neurônios no cortex (9).

O modo como os neurônios se comunicam também é muito complexo. Os neurotransmissores permitem que canais sejam abertos na membrana dos neurônios e dão origem a correntes iônicas. Existem cerca de meia centena de neurotransmissores. Dependendo do tipo dos receptores, o efeito destes neurotransmissores é excitar o neurônio por despolarizar sua membrana enquanto outros agem como inibidores por hiperpolarizarem sua membrana (6).

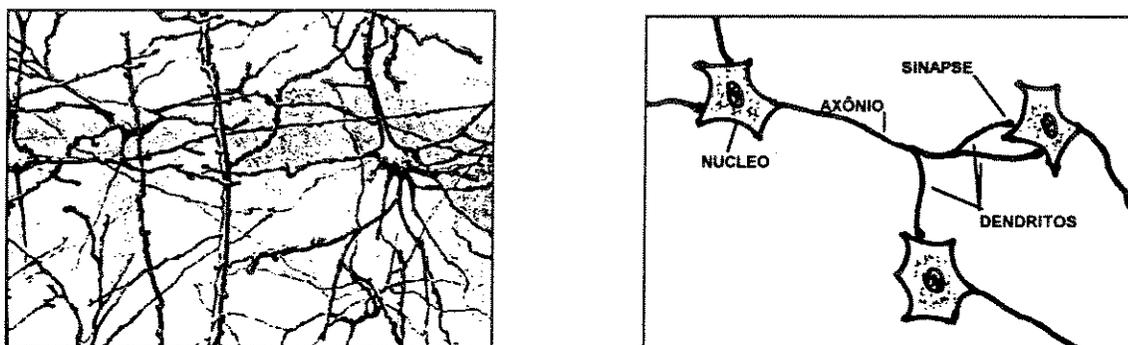


Figura 1 - Neurônios

Como neurotransmissores mais comuns temos: a Acetilcolina, a qual geralmente provoca excitação, e o ácido γ -aminobutírico, responsável pela inibição do neurônio.

Existe um grande conjunto de observações experimentais sugerindo uma grande plasticidade nas redes neurais (10,11). Assim, o cérebro apresenta um alto nível de especificidade em arquitetura e conectividade e ao mesmo tempo um alto grau de plasticidade em sua habilidade de reorganização e regeneração dos novos padrões de conectividade.

O ponto de contato entre um terminal do axônio e o dendrito de outro é denominado de sinapse (do grego *synapsis* = ação de juntar). É pela sinapse que os neurônios se unem formando as redes neurais. A intensidade do sinal transmitido depende da intensidade da junção sináptica. Esta junção sináptica sofrerá modificação durante o processo de aprendizagem e sob este aspecto a sinapse pode ser considerada como a unidade básica de memória do cérebro.

Para que possam desempenhar suas funções, o neurônio possui uma diferença de potencial entre seu interior e exterior denominado de potencial de membrana. Quando o neurônio está em repouso o potencial de membrana mantém-se constante e igual a -60mV , o que significa que a voltagem do interior do neurônio é negativa em relação ao meio extra celular. Este valor é denominado potencial de repouso. Se a voltagem no interior da membrana do neurônio tornar-se maior que o potencial de repouso, diz-se que ocorreu uma despolarização, quando ocorre o oposto, temos uma polarização.

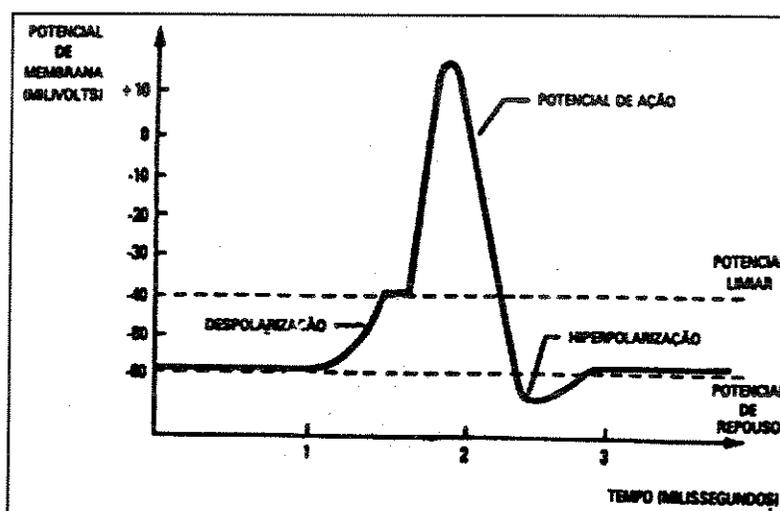


Figura 2 - Fases do Potencial de Ação

Por exemplo, se o neurotransmissor Acetilcolina entra em contato com a membrana do neurônio, abrem-se canais que permitem a entrada de íons de sódio Na^+ do meio extra celular para o interior da membrana e a saída de íons de potássio K^- . Então, a célula despolariza-se e ocorre um potencial de ação no qual este desequilíbrio (entrada do Na^+ saída de K^- e despolarização) atinge aproximadamente até +10mV (FIGURA 2) antes dos canais serem fechados, o sódio retornar ao exterior, e o potássio ao interior através da Bomba de Sódio/Potássio restaurando assim, o equilíbrio. Todo este processo tem duração aproximada de um milissegundo.

Se a estimulação for suficientemente intensa e prolongada, obtemos um trem de impulsos; uma seqüência de potenciais de ação, numa freqüência que depende do estímulo.

No corpo celular de um neurônio, chegam através das sinapses, potenciais de ação de vários outros neurônios. Seja J_j a intensidade da sinapse e S_j o potencial de ação que chega vindo de outro neurônio. Então, a contribuição de N neurônios através das sinapses será:

$$h_i = \sum_{j=1}^N J_{ij} S_j \quad (2.1)$$

Se este potencial for maior que um determinado limiar (~40mV), o neurônio irá despolarizar-se e desferir um potencial de ação através de seu axônio que irá excitar ou inibir outros neurônios ligados ao axônio deste por sinapses.

Em 1952, Hodgkin e Huxley (12) formularam um conjunto de equações matemáticas, baseado em suas observações de geração e propagação dos potenciais de ação. Essas equações, são bastante complexas e quando resolvidas numericamente, dão origem a um tipo de estudo de uma rede mais realista levando em consideração uma grande parte dos detalhes do neurônio já conhecidos.

Tal tipo de simulação, com neurônios mais próximos do real, têm aí sua força e fraqueza. A medida que o modelo torna-se mais real por adicionar mais variáveis e parâmetros (13), corre-se o risco que da simulação terminar tão pobremente compreendida quanto a própria rede. Igualmente preocupante, é o fato de que, por não conhecermos todos os detalhes do neurônio, algumas propriedades extremamente importantes podem inadvertidamente ficarem de fora, invalidando

assim, os resultados. Finalmente as simulações demandam um tempo computacional muito grande. Apenas recentemente foi possível, devido ao progresso dos computadores, ir além de casos simples.

Procura-se então, simplificar a representação do neurônio e ao mesmo tempo capturar seus princípios mais importantes. O estudo de modelos simplificados de redes neurais pode criar um base conceitual para uma melhor compreensão das redes neurais reais.

Esses modelos simplificados abstraem a complexidade dos neurônios individuais, e de seus padrões de conectividade para torná-los analiticamente tratáveis.

Para um modelo de rede neural (14), a unidade análoga ao neurônio real é denominada unidade de processamento. Esta unidade, possui muitas entradas (dendritos) e as combinam, somando-as. Esta soma é então modificada por uma função de transferência que deixa a informação passar se a soma atingir um determinado valor mínimo (limiar), ou ainda, ser uma função contínua dos valores combinados das entradas (FIGURA 3).

A saída de uma unidade de processamento pode então, ser ligada às entradas de outras unidades por conexões cujos valores modificam a intensidade do sinal que chega à próxima unidade antes de somá-las. Assim a função de somar, é tratada como uma soma ponderada pelos valores destas conexões (sinapses).

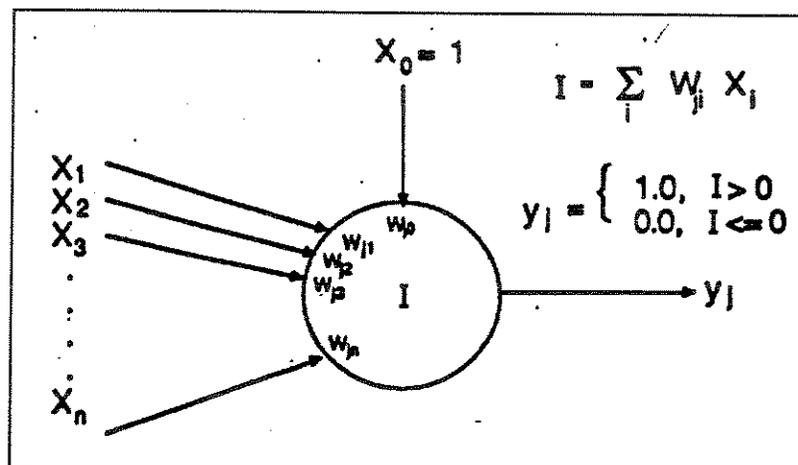


Figura 3 - Unidade de Processamento (Ref. (14))

Uma rede neural consiste portanto de muitas unidades de processamento conectadas entre si. Nesta rede, as unidades são geralmente organizadas em camadas, com conexões entre elas que podem ou não ser aleatórias. (FIGURA 4).

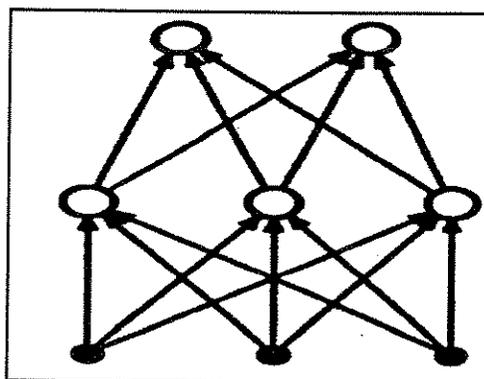


Figura 4 - Exemplo de uma Rede Neural

Existe uma camada de entrada na qual os dados são apresentados à rede e outra de saída que fornece a resposta da rede a uma determinada entrada. As outras camadas, são denominadas camadas internas da rede.

A operação de uma rede consiste em duas fases: a fase de aprendizado e a de operação.

Durante a fase de aprendizado, os valores das conexões, ou pesos, mudam, adaptando-se em resposta aos estímulos recebidos pela rede. Se desejamos uma determinada resposta da rede, comparamos sua saída a esta resposta, a fim de que, a rede possa corrigi-la. Tal tipo de aprendizado possui então um supervisor e é então denominado de aprendizado supervisionado.

Se a saída desejada é diferente da entrada, a rede é denominada de heteroassociativa. Se desejamos na saída um vetor igual ao vetor de entrada, a rede é denominada de autoassociativa. Se a saída desejada não é fornecida à rede, o aprendizado é denominado não-supervisionado.

Um terceiro tipo de aprendizado utiliza-se da regra do reforço, no qual o supervisor apenas indica se a resposta a uma entrada é boa ou ruim.

Qualquer que seja o tipo de aprendizado, a característica essencial, é a regra de aprendizado que especifica como os pesos adaptam-se em resposta a um dado

conjunto, entrada/saída. Geralmente, o aprendizado requer a apresentação do conjunto entrada/saída, muitas vezes à rede.

Na fase de operação, a rede irá responder a um dado estímulo. Normalmente esta fase é parte do processo de aprendizado, quando por exemplo a resposta desejada é comparada a saída da rede para determinarmos o erro.

A forma mais simples de reconhecimento, ocorre quando não existe retro-alimentação de uma camada para outra, ou para a mesma. Então, a informação flui de forma direta da camada de entrada para a de saída mediada pelas funções de transferências e pelas características particulares de tal tipo de rede.

Se ocorre retro-alimentação, a informação é iterada na rede até atingir algum critério de convergência. Em redes retro-alimentadas, uma estratégia de controle consiste em minimizar uma função de energia que podemos associar a cada estado da rede. O processo de operação modifica iterativamente o estado da rede conduzindo-a a um estado estável de menor energia.

A rede pode ainda operar de forma síncrona ou assíncrona. O comportamento síncrono, significa, que todas as unidades de processamento da rede mudam seus estados simultaneamente, a cada período de tempo característico da rede. Já no comportamento assíncrono, as unidades mudam seu estado de forma randômica, e independente das outras unidades da rede.

Numa rede neural, o aprendizado se dá através da apresentação de exemplos, e de regras de aprendizado que modificam as conexões sinápticas em resposta as entradas e, (opcionalmente) às saídas desejadas a estas entradas.

No aprendizado não-supervisionado, são apresentados à rede apenas os padrões de entrada e esta organiza-se internamente de forma a que cada unidade de processamento responde diferentemente a um estímulo ou grupo de estímulos.

As regras para o aprendizado não-supervisionado dividem-se em:

Regra de Hebb: a intensidade da conexão entre um caminho de entrada de uma unidade de processamento e sua saída é aumentada se ambas estão ativas.

Regra do aprendizado competitivo: neste tipo as unidades de processamento competem entre si e aquela com a resposta mais

intensa a uma dada entrada modifica suas conexões de forma a tornar-se mais semelhante à entrada.

Já no aprendizado supervisionado, a cada estímulo, a resposta desejada é também apresentada à rede e esta configura-se de forma a obter a resposta. A regra de aprendizado é conhecida como:

Regra Delta: esta variante procura reduzir o erro entre a entrada e a saída desejada de uma unidade de processamento (Backpropagation).

A memória numa rede está distribuída nos pesos das conexões entre as unidades de processamento. O valor dos pesos representa o estado de aprendizado da rede. A memória é ainda associativa no sentido de que se à rede treinada apresentamos uma entrada que contenha apenas uma informação parcial, a rede escolherá o padrão que mais se aproxima da resposta àquele apresentado na entrada.

Se a rede é autoassociativa, a apresentação de padrões incompletos à rede, resultará na sua total regeneração.

A natureza associativa e distribuída da rede neural, produz uma resposta razoável quando o estímulo é incompleto, com ruído, ou a um padrão nunca antes apresentado à rede. Esta última propriedade, constitui, um processo de generalização da rede.

Devido a natureza distribuída da memória na rede neural, observamos que a rede age como um sistema tolerante a falhas, pois, ao danificarmos ou mesmo destruímos algumas poucas unidades de processamento, o comportamento da rede sofrerá pouca alteração pois a informação não está contida num único local na rede mas sim distribuída pelas conexões de todas as unidades de processamento. Esta propriedade torna as redes neurais extremamente úteis no controle de aplicações, onde a falha no controle de um equipamento implicaria em grandes danos, tais como em reatores nucleares, usinas hidroelétricas etc....

As redes neurais, são também importantes na tarefa de reconhecimento de padrões. Tal tarefa, requer a habilidade de compararmos uma grande quantidade de

informações simultâneas, e obtermos como resposta uma generalização destes dados. É necessário, também, uma resposta razoável a padrões incompletos, ou com ruído.

A computação neural, é a que mais se aproxima da forma de percepção e reconhecimento de padrões do cérebro humano:

2.2 Principais modelos de redes neurais (1943 - 1982).

Warren McCulloch e Walter Pitts 1943.

Talvez o primeiro modelo de uma rede neural a ser estudada foi o de Warren McCulloch e Walter Pitts, (15). Sem dúvida esta foi a primeira tentativa de compreender o que uma rede neural estaria realizando, dando-se como elementos primitivos as abstrações das propriedades dos neurônios como eram conhecidas em 1943.

McCulloch e Pitts aplicaram a lógica simbólica ao problema de descrever o que as redes neurais podem fazer, eles provaram que todos os processos os quais podem ser descritos com um número finito de expressões simbólicas, como por exemplo, a aritmética, classificação, recursividade de regras lógicas etc..., podem ser incorporadas em redes do que eles denominaram de neurônios formais ou lógicos.

Os neurônios lógicos, ou do tipo McCulloch Pitts (MP) são uma representação extremamente simplificada do neurônio real. São síncronos e do tipo binário, isto é, admitem apenas dois estados. Cada neurônio possui um limiar constante podendo receber estímulos de sinapse excitatórias todas com valores idênticos. Podem ainda, receber entrada de sinapses inibitórias cuja ação é absoluta, ou seja, se a sinapse inibitória está ativa, o neurônio não dispara seu potencial de ação. Existe um intervalo de tempo para integração das entradas das sinapses.

Seu modo de operação é simples. Durante o intervalo de tempo de integração das entradas, o neurônio soma os valores dos potenciais da sinapse e se nenhuma sinapse inibitória esta ativa, e ainda se o valor encontrado for maior que o limiar então o neurônio torna-se ativo (ou dispara um potencial de ação); caso contrário ele permanece inativo.

O resultado mais importante da rede foi que qualquer expressão lógica finita pode ser realizada por um neurônio do tipo MP. Este resultado é muito importante por mostrar que elementos simples conectados entre si, formando uma rede, possuem um grande poder computacional. Uma vez que tais elementos foram baseados na neurofisiologia, este resultado sugere que o cérebro é uma máquina extremamente lógica e poderosa.

McCulloch e Pitts estavam cientes de que muitos dos fenômenos ocorrendo no sistema nervoso eram de origem contínua e não discreta e em seu trabalho, comentaram sobre a importância de que mudanças contínuas no limiar dos neurônios poderiam ter no aprendizado e adaptação.

Pitts e McCulloch abordaram também em 1947 (16) o problema de reconhecimento de padrões, problema central de qualquer teoria do comportamento inteligente. Eles observaram que os animais necessitam reconhecer diferentes versões do mesmo padrão, assim como, necessitamos reconhecer diferentes versões de um mesmo texto, letras maiúsculas ou minúsculas, cores diferentes etc. Construíram um modelo composto de duas redes neurais para resolver o problema de reconhecimento de padrões. A primeira rede procura encontrar propriedades invariantes de um dado padrão, isto é, propriedades comuns a todas as possíveis variantes do padrão. A segunda rede transforma qualquer variante apresentada no padrão modelo.

Donald Hebb 1949.

O próximo avanço vem a partir da publicação do livro "The Organization of Behavior" de Donald Hebb (17) em 1949, que foi o primeiro a estabelecer que o aprendizado se dá via modificação nas sinapses entre os neurônios.

A regra, hoje conhecida como regra de Hebb, estabelece que **"quando um axônio da célula A está próximo bastante para excitar uma célula B e repetidamente ou persistentemente toma parte em estimulá-la, algum processo de crescimento ou mudança metabólica ocorre em uma ou ambas as células tal que a eficiência de A como uma das células que estimula B aumenta"**.

Hebb ainda propõe que devido às modificações sinápticas, ocorre a formação de agrupamentos de células que possuem um comportamento síncrono. Seguiu uma antiga sugestão de Cajal (3) e postulou que a ativação repetida de um neurônio por

outro através de uma dada sinapse aumenta sua condutância de forma a que grupos de células fracamente conectadas, se ativadas de forma síncrona tendem a organizar-se em conjuntos mais fortemente conectados. Esta talvez, seja uma das principais contribuições de seu trabalho pois, embora houvesse na época falta de evidências para o suporte de suas idéias, a teoria dos conjuntos síncronos de células despertou uma enorme investigação de como esta atividade neural síncrona é gerada e propagada nas redes neurais.

Sua idéia de que existem padrões temporais estáveis de atividade neural reapareceu posteriormente em vários modelos de "atratores" para a atividade neural (44,49).

Hebb ao expor suas idéias tinha conhecimento da natureza distribuída que o sistema nervoso utiliza e estava a par das idéias de Karl Lashley sugerindo uma representação distribuída.

Karl Lashley 1950.

Lashley (18) descobriu que ao lesionar várias partes do cérebro de um rato, o animal mesmo assim, era capaz de demonstrar capacidade de aprendizado.

Esses dados levaram Lashley a concluir que existe relativamente pouca localização no cortex cerebral. Técnicas modernas mostram que as conclusões de Lashley não são verdadeiras e que existe uma grande localização de funções no cortex como por exemplo, no mapeamento de frequências na audição, no mapeamento de espaço visual na superfície do cortex, etc. Este mapeamento é, contudo, passível de re-organização em alguns casos de lesão, o que provocou as conclusões de Lashley.

Entretanto, é igualmente errado supor que a informação armazenada está localizada apenas num pequeno conjunto de células e suas interconexões. Por exemplo, uma imagem visual excita um grande número de neurônios do sistema visual; exatamente quantos é uma questão em debate (19,20).

Podemos então supor que a informação é distribuída e que muitos elementos devem estar simultaneamente ativos para representar a informação.

As conclusões de Lashley favorecem mais um sistema distribuído que localizado.

Brian G. Cragg e H. N. V. Temperley (1954).

Em 1954 o físico Temperley e o neuroanatomista Cragg (21) observaram que assim como os neurônios podem estar ativados (emitindo um potencial de ação) ou em repouso, átomos numa rede que possuem momento magnético podem estar num dos dois estados de energia, por exemplo, com o spin apontando para cima ou para baixo. Assim, como os neurônios excitam ou inibem seus vizinhos, também os átomos exercem forças uns sobre os outros tendendo alinharem seus spins na mesma direção ou direção oposta. Desta forma, as propriedades dos neurônios numa rede densamente conectada, são similares à uma rede de átomos com spins.

Átomos num material podem apresentar diferentes tipos de ordem (ou desordem). Estes tipos vão desde ordem de curto alcance na qual, em média, cada spin apontando para cima é rodeado apenas por spins apontando na direção oposta, como por ordem de longo alcance na qual em média, spins apontando numa direção persistem por uma distância maior. Tais tipos de modelos nos dão uma boa explicação das propriedades dos materiais magnéticos. Por exemplo, num material ferromagnético, os spins tendem a forçarem seus vizinhos a orientarem-se na mesma direção, daí uma ordem de longo alcance.

Já num material antiferromagnético, os spins tendem a forçarem seus vizinhos na direção oposta e possuem também ordem de longo alcance.

Seria de esperar que as redes neurais exibissem um comportamento "análogo". Cragg e Temperley sugeriram então que:

Os padrões apresentados pelos materiais ferromagnéticos compreendendo spins apontando para cima e para baixo seriam análogos às regiões das redes neurais que possuem neurônios disparando potenciais de ação e neurônios em repouso.

Que as redes neurais poderiam apresentar efeitos similares à da histerese ferromagnética nas transições entre estados desordenados para ordenados. Isto implicaria que padrões apresentados pelas redes neurais uma vez disparados por

estímulos externos, seriam estáveis contra uma atividade randômica espontânea e poderiam desta forma constituir "uma memória" do estímulo (22).

Este trabalho daria origem ao trabalho de Little (38) 20 anos depois.

B. Farley e W. A. Clark 1954.

A. M. Uttley 1954.

Perto de 1950, entrou-se na era da simulação por computador. Quando é proposta uma regra de aprendizado, é então, possível simulá-la para observar se funciona. Uma das primeiras simulações foi feita por Farley e Clark em 1954 (23), que estudaram alguns detalhes das propriedades de neurônios interconectados.

Uttley em 1954 (24) demonstrou que redes neurais com conexões sinápticas modificáveis do tipo Hebb, poderiam aprender a classificar conjuntos simples de padrões binários do tipo 110101101, 100011001 etc em classes equivalentes, por exemplo, todas as que iniciam com 100.

Nathaniel Rochester e outros 1956.

A simulação mais interessante foi feita por Nathaniel Rochester e outros em 1956 (25).

Em 1956, Marvin Minsky, John McCarthy, Nathaniel Rochester e Claude Shannon (14) organizaram a primeira conferência sobre inteligência artificial promovida pela Fundação Rockefeller. Nesta reunião, pesquisadores de todo mundo uniram-se para discutir o uso dos computadores na simulação de "cada aspecto do aprendizado ou qualquer outra propriedade da inteligência". Foi nessa conferência que as áreas da computação neural e inteligência artificial foram iniciadas e que Rochester, da IBM Research apresentou um modelo de rede neural construído com várias centenas de neurônios.

Rochester simulou então, em seu trabalho, a regra de aprendizado de Hebb. Sem dúvida foi a primeira tentativa de testar a teoria de uma rede neural bem formulada através de uma simulação no computador.

O aspecto mais interessante dessa simulação foi o de fazer uma série de hipóteses que não estavam presentes na teoria original de Hebb. Simulações envolvem uma grande precisão de detalhes que não podem ser ignorados ou apenas discutidos de forma qualitativa.

Os neurônios eram do tipo binário com um limiar. Os autores descobriram que a medida que o sistema evoluía no tempo, a intensidade da sinapse do tipo Hebb, aumentava sem limite. Para resolver este problema propuseram que a soma de todos os valores das sinapses fosse constante. Foi necessário ainda a introdução de um fator de fadiga onde células que já dispararam seu potencial de ação são menos prováveis de novamente dispará-los num futuro próximo.

Quando o sistema foi simulado, não mostrou o aparecimento dos grupos síncronos de células predito por Hebb, ou seja, não mostrou a auto-organização esperada. Na segunda tentativa foi incorporada a inibição, no sentido de que um grupo de células inibe em sua vizinhança outros de tornarem-se ativos. As sinapses assumiam valores entre +1 e -1 ao invés de +1 e 0. O neurônio foi também modificado, sua saída era relacionada à frequência de disparos que podiam assumir valores entre 0 e 15 e não mais unicamente ± 1 .

A regra de Hebb foi modificada de tal forma que se as frequências das atividades pré e pós sinápticas eram correlacionadas, a intensidade sináptica aumentava e ainda, se um neurônio disparava um potencial de ação enquanto o outro estava inativo, a intensidade decrescia. Esta é uma generalização natural da regra de Hebb que é comumente usada até hoje.

A atividade da célula pós sináptica foi obtida por somar a atividade pré-sináptica mediada pela intensidade da sinapse, regra utilizada ainda nos dias de hoje. Usou-se também a hipótese de que os neurônios conectam-se apenas a seus vizinhos mais próximos.

Usaram na simulação cerca de 512 neurônios. Assumindo os dados acima, a simulação funcionou, e ocorreu a formação dos grupos síncronos de células com conexões excitatórias, entre as células destes grupos, e inibitórias entre grupos diferentes.

Frank Rosenblatt 1958 (perceptrons).

O próximo passo foi dado por Frank Rosenblatt em 1958 (26), que mostrou como redes neurais com neurônios do tipo lógico e com conexões modificáveis poderiam ser "treinadas" para classificar certos conjuntos de padrões como similares ou distintos. Rosenblatt denominou tais redes de "perceptrons".

O perceptron é um sistema classificador de padrões; o primeiro perceptron era capaz de algum aprendizado e possuía um alto grau de plasticidade. Era também, capaz de fazer generalizações limitadas e podia classificar padrões com ruído.

O perceptron (FIGURA 5) tinha inicialmente o propósito de reconhecimento óptico de padrões. Uma rede de 400 fotocélulas correspondente aos neurônios sensíveis à luz da retina recebiam o estímulo óptico inicial. Estas fotocélulas eram conectadas a unidades associadoras que recolhiam os impulsos elétricos das fotocélulas. Conexões entre as fotocélulas e as unidades associadoras eram feitas de forma randômica. Se a entrada de alguma fotocélula excedia um certo limiar as unidades associadoras produziam uma resposta.

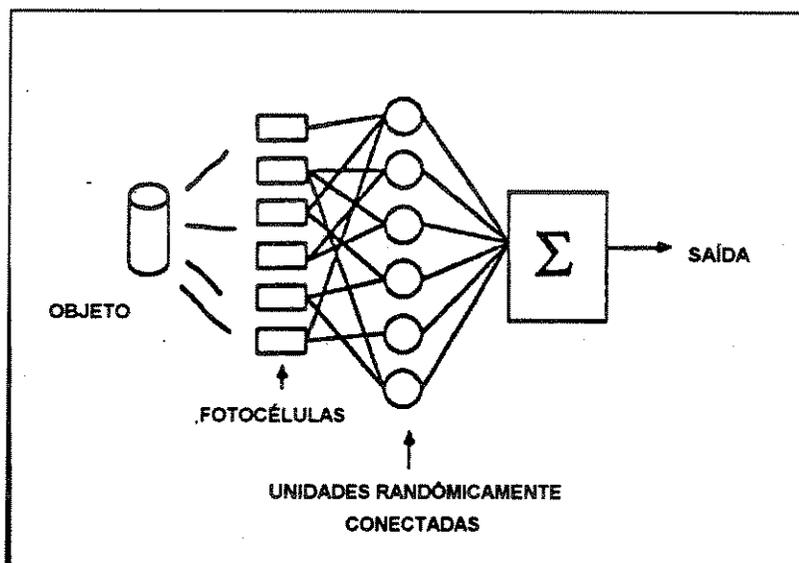


Figura 5 - Perceptron - (Ref. (14))

Uma das características importantes dos perceptrons provada por Roseblatt, é que dado um conjunto de entradas e saídas ± 1 , existe uma forma de treinar os pesos que garante as saídas desejadas se e apenas se um conjunto de pesos existir.

Minsky e Papert (27) provaram que para os conjuntos "mais interessantes" não existe nenhum conjunto de pesos. Mostraram ainda, que um conjunto de pesos existe se o conjunto de entrada é linearmente separável e, a partir daí, demonstraram que o perceptron não seria capaz de efetuar uma operação do tipo ou exclusivo.

Conjunto linearmente separável significa que, dado um espaço de todas as entradas, todas aquelas que possuem uma determinada propriedade devem estar de um único lado de um hiperplano e todas as outras com a outra propriedade do lado oposto.

Minsky e Papert destacaram ainda, que embora um perceptron multi-camada poderia estar apto a resolver o problema do tipo ou exclusivo, não existia forma de responder quais pesos deveriam ser ajustados quando a saída fosse indesejada.

Então, baseados na inabilidade de um perceptron simples em resolver "problemas interessantes", concluíram que a pesquisa de máquinas semelhantes seria de pouco valor.

O fato do perceptron ser capaz de um comportamento adaptativo complexo, tornou-o atrativo para os engenheiros da época que vislumbraram várias aplicações para seu uso. Seu comportamento complexo foi difícil de ser analisado matematicamente (28,29) e tal fato conduziu a várias conclusões sobre o poder e as limitações do aprendizado pelas máquinas.

O importante a ser observado é que o modo apropriado para o estudo do cérebro seria vê-lo como um associador para o aprendizado e que a linguagem da lógica simbólica ou Booleana não era o melhor caminho.

O ponto principal era de que num ambiente cheio de ruído e variável, e com conexões sinápticas não completamente pré-especificadas, seria então necessário ao sistema, aprender a fazer as conexões baseados no estímulo. Era o início do "errando é que se aprende".

Oliver Selfridge 1958.

Em 1958, Oliver Selfridge (30) propõe o Pandemonium que engloba técnicas de otimização, no qual procura-se maximizar uma determinada função do estado do sistema.

Através deste esquema de aprendizado, a resposta de um dado neurônio depende de suas entradas, isto é, da intensidade das conexões sinápticas entre a unidade e o padrão de entrada ou entre a unidade e as camadas anteriores. Se conhecermos os padrões de entrada podemos modificar estas conexões até maximizarmos a efetividade do sistema. Selfridge sugere que podemos fazer isto por modificar os pesos um pouco em todas as direções e escolher a direção em que se dá um maior aumento na efetividade e assim sucessivamente. Isto assegura que a resposta do sistema será aumentada ou, na pior das hipóteses será não decrescente.

Bernard Widrow e Marcian Hoff 1960

Logo após os perceptrons, Bernard Widrow e Marcian Hoff em 1960 (31-33), propuseram o ADALINE (Adaptative Linear Neuron) cuja saída era ± 1 em vez do 0 e 1 dos perceptrons (FIGURA 6).

O ADALINE consiste de uma série de pesos adaptáveis. Uma entrada constante de +1 é sempre conectada ao ADALINE. As entradas assumem somente valores ± 1 . Os pesos podem ser positivos ou negativos. A soma ponderada das entradas incluindo a entrada constante +1, é aplicada à função de transferência, a qual converte a saída para valores ± 1 .

Um mecanismo de controle do aprendizado testa a entrada, a saída, a saída desejada e usa esta informação para ajustar os pesos. Inicialmente, os pesos adotam valores aleatórios.

O algoritmo de ajuste dos pesos do ADALINE é o seguinte:

Aplicando-se o padrão a ser classificado na entrada do ADALINE, calcula-se o erro (a diferença entre a soma ponderada e a resposta desejada). Ajustam-se os pesos, a fim de reduzir o erro. Este processo é repetido para o próximo conjunto de entrada.

Calcular o erro para cada iteração e ajustar os pesos para diminuir ou eliminar o erro é no que consiste a regra de aprendizado do Widrow-Hoff. Eles assumiram ainda a presença de um supervisor conhecendo qual a resposta a um dado estímulo.

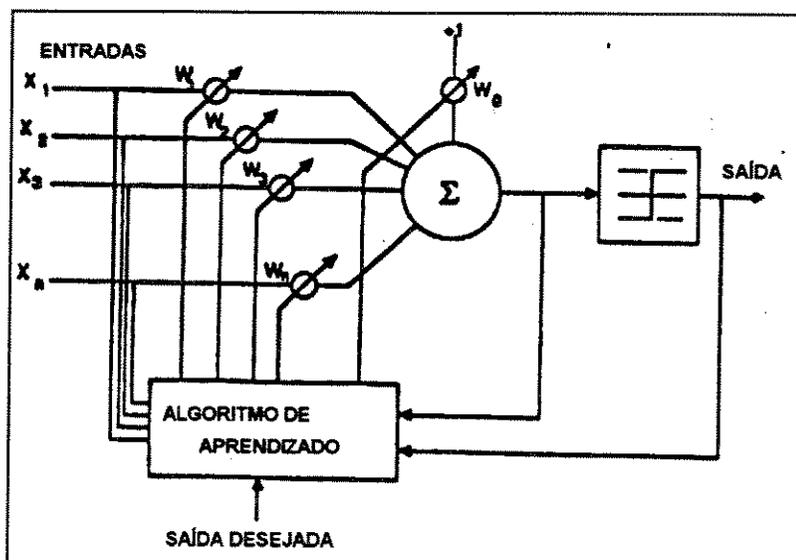


Figura 8 - Adaline - (Ref. (14))

Sabia-se também a resposta efetivamente dada. Assumiram então, que o sistema poderia calcular o erro entre o padrão conhecido e a resposta efetivamente dada, ou seja a diferença entre eles. Os pesos sinápticos são então ajustados de forma a reduzir o erro até atingir o valor zero.

O algoritmo garante a convergência se existir um conjunto de pesos, e este conjunto reduz o erro no sentido dos mínimos-quadrados. Este algoritmo é também conhecido como LMS (Least Mean Square) em processamento de sinais (34).

Um aspecto importante da idéia de correção de erro é a existência de um supervisor que possui um perfeito conhecimento da resposta. Tais algoritmos são conhecidos como supervisionados. O algoritmo "backpropagation" desenvolvido recentemente é o mais conhecido tipo de algoritmo supervisionado. (35).

Marvin Minsky e Seymour Papert 1969.

Até 1969 tudo corria muito bem para a área das redes neurais, centenas de algoritmos, várias regras de aprendizagem, e a pesquisa incentivada por fundos de várias instituições.

Então, com a publicação do livro "Perceptrons" de Marvin Minsky e Seymour Papert (27) iniciou-se um período ruim para a área. Por que tal livro provocou efeitos tão devastadores?

Muitas foram as causas, mas a principal foi que o livro em questão abordava limites da computação neural. Roseblatt, já estava a par dos problemas com os perceptrons: havia problemas não resolvidos pelos mesmos, havia o problema da separabilidade linear das entradas para a sua classificação. Existiam também problemas se muita generalização era requerida. Algumas destas dificuldades foram abordadas no último capítulo do livro de Roseblatt.

Minsky e Papert colocaram o estudo das limitações dos perceptrons em uma base sólida.

O principal resultado matemático do livro é sobre o predicado geométrico da conectividade. Uma figura A é conexa, se pudermos traçar uma curva contínua, totalmente contida em A através de quaisquer dois pontos de A . É possível então demonstrar que um perceptron de diâmetro limitado não consegue simular o predicado geométrico da conectividade..

Mas bastavam estes resultados técnicos para interromper toda uma linha ativa de pesquisa? A resposta é que as redes neurais estavam em declínio já por alguns anos porque, apesar do grande sucesso inicial, falharam quanto à expectativa futura. Resultados práticos não se materializaram e resultados de sucesso no novo campo da inteligência artificial acabaram por obscurecer as redes neurais.

Assim, o aparecimento do livro foi um golpe final do processo que arrastava-se já por algum tempo. No capítulo final do livro, estavam reservadas palavras que teriam grande impacto sobre a pesquisa futura. Lá os autores expressavam a crença de que as limitações que descobriram para o perceptron simples deveriam ser verdadeiras para outras variantes do perceptron, mais especificamente, para perceptrons multicamadas. Esta conjectura refreou o financiamento de verbas para pesquisas na área.

E por que versões mais complexas do perceptron deveriam sofrer das mesmas limitações?

Felizmente, para todos os que trabalham com redes neurais os autores do livro estavam errados.

Teuvo Kohonen 1972.

James A. Anderson 1972.

Em 1972, as redes associativas retornam nos trabalhos de Teuvo Kohonen (36) e Anderson (37) que de forma independente obtiveram os mesmos resultados.

No modelo denominado associador linear, os neurônios eram do tipo analógico e não binário. Suas entradas eram então multiplicadas pelos pesos sinápticos e somados, sendo sua saída proporcional a esta soma. Neste modelo, dado um vetor de entrada, o vetor de saída é calculado pela matriz de conexão vezes o vetor de entrada. A relação entrada/saída é então especificada por uma simples multiplicação de uma matriz por um vetor, uma transformação linear.

A regra de aprendizado era uma generalização da regra de Hebb, na qual a mudança sináptica era proporcional ao produto das atividades pré e pós sináptica (Anderson), ou à correlação entre os elementos do vetor de entrada e do vetor de saída (Kohonen), que são a mesma coisa.

A matriz de conexão era então calculada pelo produto externo entre os vetores de entrada e saída.

Se o número de entradas e saídas é grande, haverá múltiplas correlações entre os diferentes elementos no nível sináptico. Esta propriedade fortalece a habilidade da rede em memorizar padrões. Ela também remove a assimetria entre entrada e saída. Se ambos são vetores, é então fácil realimentar a entrada com a saída. Tal tipo de rede, Kohonen denominou-a de autoassociativa e será a base das redes do tipo Hopfield. Se os vetores de entrada e saída são diferentes a rede é denominada de heteroassociativa.

O único caso em que há associação perfeita é quando os vetores de entrada são ortogonais. Este fato, coloca um limite no número de vetores que podem ser armazenados. Simulações futuras indicariam que o sistema trabalha bem para vetores aleatórios cujo número máximo de vetores armazenados está situado entre 10% e 20% do número total de neurônios.

Willian A. Little 1974.

Vinte anos após o trabalho de Gragg e Temperley (21,22), Willian A. Little (38) retomou a analogia entre uma rede neural e um sistema de spins interagindo entre si.

A existência de estados persistentes numa rede neural corresponde à ocorrência de ordem de longo alcance no sistema de spins.

Little considera uma rede na qual os neurônios estão isolados de um estímulo externo, e na qual disparam potenciais de ação de modo síncrono, isto é disparam em um tempo que é múltiplo inteiro de um período T da ordem do período refratário do neurônio. Em seu trabalho Little propõe um modelo probabilístico para o neurônio em vez de um modelo determinístico.

Assim, é possível calcular a probabilidade de passarmos de um estado a outro dado os pesos sinápticos. Se a rede possui N neurônios, cada um dos quais pode estar no estado ± 1 , existem então 2^N possíveis estados para o sistema. Consideremos uma matriz ($2^N \times 2^N$) de probabilidades P com todas as probabilidades de transições entre os diferentes estados. Um elemento desta matriz $P_{\alpha\beta}$ representa a probabilidade de transição do estado α para o estado β . Um estado será persistente no sentido de que a transição para ele próprio, isto é o termo diagonal $P_{\alpha\alpha}$ será próximo de 1. Trata-se de um modelo de Markov, onde a probabilidade de transição só depende do estado corrente.

O estado da rede é descrito por:

$$\psi(t) = |S_1, S_2, \dots, S_N\rangle \quad (2.2)$$

onde: $S_i = +1$ se o neurônio disparou um potencial de ação e $S_i = -1$ caso contrário.

Esta configuração determina o estado da rede neural num instante particular e corresponde à configuração dos spins numa linha de uma rede de átomos com spins.

A configuração que descreve o estado da rede neural após o próximo ciclo corresponde à configuração dos spins da próxima linha. Assim, a existência de correlação entre dois estados da rede neural separados por um longo período de tempo corresponde à correlação espacial de um sistema de spins de Ising bidimensional.

Como a existência de ordem de longo alcance em um sistema de spins é formalmente idêntica ao aparecimento de estados persistentes da rede neural, Little, usou os resultados dos sistemas de spins para demonstrar que a existência de ordem de longo alcance está associada à degenerescência do maior autovalor de matriz de transferência.

Little, mostrou ainda, que uma transição para um estado persistente é análoga a uma transição de fase.

Num trabalho subsequente Little e Shaw (1975) (39), estenderam o trabalho anterior de Little e propuseram um modelo no qual os neurônios possuem apenas dois estados, (disparando ou não um potencial de ação) mas, com limiar variável. Basearam o modelo, na observação de que havia, no neurônio real, muitas fontes de ruído. Existe por exemplo, um processo de Poisson governando o número de "quanta" (vesículas sinápticas) eliminados quando a sinapse é ativada, bem como uma descarga espontânea dos mesmos.

Assim dada uma certa despolarização a probabilidade de um disparo ocorrer é uma função tanto do limiar da célula quanto da sua variabilidade interna.

A célula pode disparar um potencial de ação mesmo quando a despolarização estiver abaixo do valor limiar, ou permanecer em repouso mesmo quando estiver pouco acima deste valor. O conceito de um neurônio probabilístico teria aplicações futuras (40).

D. Marr e T.Poggio 1976.

Em 1976 D. Marr e T.Poggio (41), estudaram como o cérebro humano constrói a noção de profundidade que nos permite a visão tridimensional.

Trata-se de um problema de um algoritmo cooperativo no qual um número de elementos observando apenas uma parte do problema interage com a outra parte e converge para uma solução global comum.

Marr e Poggio, através de seus algoritmos criaram uma matriz de conexões à qual era apresentada uma imagem (estímulo externo). Iniciava-se um processo iterativo no qual as saídas de cada célula era 0 ou 1 e atingia-se um estado estável que ao ser iterado novamente, não sofria mais mudanças.

A dinâmica deste processo é semelhante ao modelo do Hopfield.

No modelo de Marr e Poggio observamos um vetor inicial representando um estado sendo operado iterativamente de acordo com regras simples e atingindo um estado final estável.

Este tipo de sistema tornou-se muito comum atualmente. Observamos ainda que sua matriz de conexão é deduzida da estrutura computacional do problema. Um exemplo deste tipo de sistema, foi desenvolvido por Hopfield e Tank (42) que usaram um rede neural para estimar a solução do problema do caixeiro viajante.

Em outros sistemas, a matriz de conexão, é formada pela utilização de uma regra de aprendizado.

Sun-Ichi Amari 1977.

Em 1977 Sun-Ichi Amari (43) propõe dois modelos. No primeiro, denominado associador de padrões um padrão x é transformado num padrão de saída y usando uma transformação T , ou seja, $y = Tx$. Amari assume que este tipo de rede possui, somente durante o processo de aprendizado, uma outra entrada, que faria o papel de supervisor até que a relação entrada-saída seja corretamente formada. O segundo, é uma rede autoassociativa, a qual Amari denomina de "rede formadora de conceito" na qual o padrão de saída é realimentada novamente à entrada da rede. Amari define como "conceito", um estado estável dessa rede, isto é, estados do tipo $x = Tx$.

Amari define uma função R dos pesos, do supervisor e das entradas, e uma função L unicamente dos pesos cuja forma é análoga à função energia que é de grande importância no desenvolvimento futuro das redes neurais. Amari menciona ainda que a solução de "energia mínima" é encontrada quando os pesos estão na mesma direção do autovetor com o máximo autovalor da matriz de covariância formada pelos padrões de entrada.

Nessas redes Amari usa a notação de que as entradas são dadas por x_1, \dots, x_N e as saídas y_1, \dots, y_M . Se w_{ij} é o valor da conexão entre o elemento i e j e h_i o limiar do i -ésimo elemento, então:

$$y_i = \varphi\left(\sum_{j=1}^N w_{ij}x_j - h_i\right) \quad i = 1, 2, \dots, m \quad (2.3)$$

ou em notação matricial

$$\bar{y} = \varphi(W\bar{x} - \bar{h}) \quad (2.4)$$

Se a rede for do tipo recorrente e do tipo "formadora de conceitos", então a saída no tempo $t+1$ é dada por:

$$\bar{x}(t+1) = \varphi(W\bar{x} - \bar{h}) \quad (2.5)$$

Um estado que satisfaz $T\bar{x} = \bar{x}$ é denominado um estado de equilíbrio da rede. Estes estados representam informação que a rede pode mantê-la (análogo aos estados persistentes de Little) indefinidamente. O estado de equilíbrio é visto por Amari como um padrão "conceito" que a rede formou.

James A. Anderson e outros 1977.

Ainda em 1977, Anderson, Silverstein, Ritz e Jones (44) propõem uma generalização do associador linear (FIGURA 7) na qual dois grupos α e β de N neurônios cada um, onde cada neurônio do grupo α está ligado a cada neurônio do grupo β . Este modelo já havia sido abordado por Leon Cooper em 1973 (45) e por Nass e Cooper em 1975 (46).

Vamos supor que os neurônios atuem como somadores lineares analógicos de suas entradas (47). Se assumirmos que a intensidade da conexão sináptica a_{ij} que liga os neurônios j no grupo α com o i -ésimo neurônio do grupo β e se $f(j)$ é a atividade do neurônio j no grupo α e $g(i)$ a atividade do neurônio i no grupo β então:

$$g(i) = \sum_{j=1}^N a_{ij} f(j) \quad (2.6)$$

Seja agora a seguinte situação:

O conjunto de neurônios α apresenta um padrão de atividades \vec{f}_1 e o conjunto β um padrão de atividades \vec{g}_1 .

Vamos associar \vec{f}_1 com \vec{g}_1 supondo que os elementos da matriz de conexão sejam inicialmente nulos. A principal hipótese é a seguinte: "para associarmos o padrão \vec{f}_1 em α com \vec{g}_1 em β necessitamos mudar o conjunto dos pesos sinápticos de acordo com o produto da atividade pré-sináptica com a atividade pós-sináptica".

Seja por conveniência

$$\|\vec{f}\| = 1 \quad \text{onde} \quad \|\vec{f}\| = \left[\sum_{i=1}^N f(i)^2 \right]^{1/2} \quad (2.7)$$

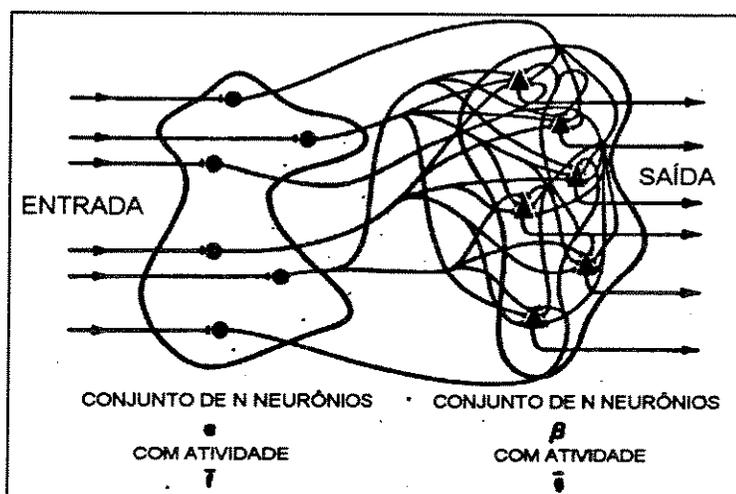


Figura 7 - Associador Linear - (Ref. (44))

A mudança na ij -ésima sinapse é dada por $f_1(j)g_1(i)$, e o conjunto das conexões forma uma matriz A_1 descrita a seguir:

$$A_1 = \vec{g}_1 \vec{f}_1^T \quad (2.8)$$

onde \vec{f}_1^T é o transposto do vetor \vec{f}_1 .

Logo

$$A_1 \vec{f}_1 = \vec{g}_1 \vec{f}_1^T \vec{f}_1 = \vec{g}_1 \quad (2.9)$$

e assim, o vetor representando o padrão \bar{g}_i aparece no grupo β o que corresponde à definição de associação.

Se tivéssemos k conjuntos de associações $(\bar{f}_1, \bar{g}_1), (\bar{f}_2, \bar{g}_2), \dots, (\bar{f}_k, \bar{g}_k)$, cada conjunto daria origem a uma matriz A_k . Então desde que assumimos que uma única sinapse participa do armazenamento de vários padrões (memória distribuída) podemos obter uma única matriz A para todos os conjuntos expressa por:

$$A = \sum_k A_k \quad (2.10)$$

Se os padrões \bar{f}_i forem mutuamente ortogonais, isto é:

$$\bar{f}_i \cdot \bar{f}_j = 0 \quad i \neq j \quad (2.11)$$

então:

$$A\bar{f}_j = \sum_k A_k \bar{f}_j = A_j \bar{f}_j + \sum_{k \neq j} A_k \bar{f}_j \quad (2.12)$$

$$= \bar{g}_j + \sum_{k \neq j} g_k \bar{f}_k \cdot \bar{f}_j = \bar{g}_j \quad (2.13)$$

Se os vetores \bar{f}_i não forem ortogonais, o efeito será a produção de ruído que poderá afetar a associação dos padrões. Se os padrões forem ortogonais, não haverá interação entre eles, e assim, a resposta de um sistema a uma entrada não será influenciada pela outra entrada.

Anderson, analisa o caso de um sistema associativo (FIGURA 8) em que o conjunto α dos neurônios é ligado a si próprio em vez de ao conjunto β .

O aumento da intensidade sináptica Δa_{ij} é proporcional ao produto da atividade do i-ésimo neurônio $f(i)$ e o j-ésimo neurônio $f(j)$.

Observamos ainda que :

$$\Delta a_{ij} = \Delta a_{ji} \quad (2.14)$$

isto é, a matriz é simétrica.

O fato de A ser simétrica implica na existência de N vetores mutuamente ortogonais $\vec{e}_1, \vec{e}_2, \dots, \vec{e}_N$ tal que:

$$A\vec{e}_i = \lambda_i\vec{e}_i \quad i=1,2,\dots,N \quad (2.15)$$

onde λ_i é um número real (48)

Os \vec{e}_i são denominados de autovetores da matriz A e os λ_i são os autovalores associados aos autovetores \vec{e}_i .

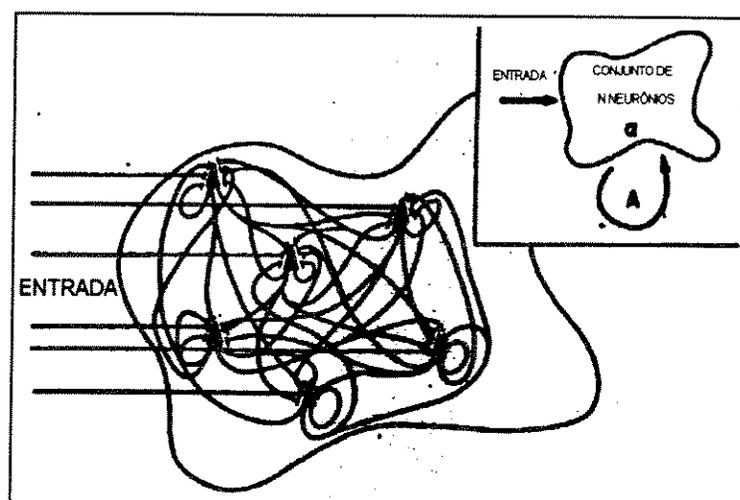


Figura 8 - Autoassociador - (Ref. (44))

Uma vez que existem N autovetores mutuamente ortogonais, cada vetor é uma combinação linear dos autovetores. Como consequência, a matriz A é completamente determinada pelo seu conjunto de autovetores e correspondentes autovalores.

Em seu trabalho, Anderson procurou o que os psicólogos denominam de características (features), que seriam uma espécie de "átomos da percepção", a partir dos quais as percepções mais complicadas são construídas. Estas características devem provavelmente serem distribuídas como padrões de descarga de muitas células. Anderson, propõe combinar o associador linear com as conexões recorrentes existentes no cortex para a construção de um sistema auto associativo.

A matriz de conexão de tal modelo, deveria possuir autovetores ortogonais que poderiam descrever as propriedades esperadas das características, pois sendo ortogonais, poderiam decompor uma característica complexa em suas componentes mais simples.

E finalmente chegamos ao modelo de Hopfield que abordaremos no próximo capítulo.

Capítulo 3

MODELO DE HOPFIELD

3.1 O trabalho de Hopfield (49).

Embora muitas das idéias de Hopfield tivessem precursores, a forma de uní-las, o tratamento matemático adequado, a introdução do conceito de que os atratores eram pontos de mínimo de uma função energia claramente definida, e a relação desta rede com o modelo de Ising, tornam o trabalho de Hopfield um dos mais importantes da área por esclarecer parte da dúvida em relação ao processo dinâmico das redes neurais.

Hopfield, em seu trabalho, inicia por destacar a emergência de propriedades de conjunto, isto é, propriedades que surgem quando conectamos um grupo de neurônios entre si e dos quais conhecemos apenas o comportamento. Surgem então, certas propriedades espontâneas que são insensíveis aos detalhes dos modelos, e que estão ausentes no comportamento individual.

A partir daí, Hopfield afirma que a função do sistema nervoso é desenvolver um número de pontos localmente estáveis no espaço de estados deste sistema que agirão como atratores para pontos próximos a eles.

Seja um sistema físico descrito por coordenadas x_1, x_2, \dots, x_N , componentes de um vetor de estado X . Sejam ainda X_a, X_b, \dots pontos limites localmente estáveis do sistema. Se o ponto inicial for $X = X_a + \Delta$ o sistema, por sua dinâmica, dirigir-se-á para o

ponto final $X \approx X_a$. Se pensarmos que os pontos X_a, X_b, \dots representam informação armazenada no sistema, observamos que o sistema age como uma memória, e o ponto inicial $X_a + \Delta$ representa o conhecimento parcial de um determinado item X_a . O sistema é então uma memória de conteúdo endereçável e sua utilidade será maior se houver uma regra de como tornar o conjunto de vetores iniciais X_a, X_b, \dots , estados estáveis do sistema.

Os neurônios da rede de Hopfield são do tipo lógico, ou seja, do tipo MacCulloch-Pitts com valores +1 quando disparando um potencial de ação e, zero caso contrário. A conexão entre um neurônio i e outro j é dada por J_{ij} . O estado de um sistema de N neurônios é determinado por especificar seus N componentes V_i , e pode também ser representado por um número binário de N bits.

A dinâmica da evolução de um estado da rede é dada por: cada neurônio de forma assíncrona, soma suas entradas sinápticas e compara ao seu limiar θ_i . Se essa soma for maior, muda seu estado para +1, se menor altera seu estado para 0, isto é

$$V_i(t+1) = H\left(\sum_{j=1}^N T_{ij} V_j - \theta_i\right) \quad (3.1)$$

onde $H(x)$ é a função de Heaviside definida por:

$$H(x) = \{1 \text{ se } x > 0 \text{ e } 0 \text{ se } x < 0\} \quad (3.2)$$

A regra de aprendizado da rede de Hopfield é a mesma de Anderson (37) e Kohonen (36) e constitui uma generalização da regra de Hebb, isto é, a mudança sináptica é proporcional ao produto das atividades pré e pós sináptica ou à correlação entre os elementos do vetor de entrada e saída.

Assim:

$$T_{ij} = \sum_{\mu=1}^p (2V_i^\mu - 1)(2V_j^\mu - 1) \quad (3.3)$$

com $T_{ii} = 0$.

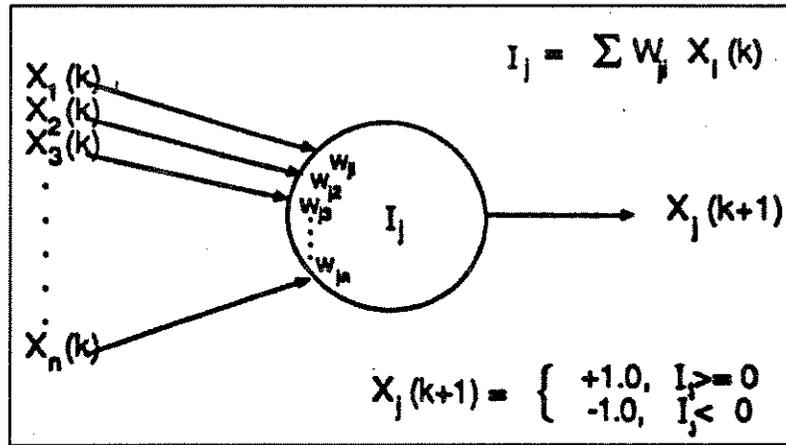


Figura 9 - Unidade de Processamento da Rede de Hopfield - (Ref. (14))

Se fizermos a transformação $\xi_i = 2V_i - 1$ observamos que os neurônios assumem os valores ± 1 em vez de 0 e 1 e a regra aprendizado fica:

$$J_{ij} = \sum_{\mu=1}^p \xi_i^{\mu} \xi_j^{\mu} \quad (3.4)$$

e a dinâmica da rede

$$S_i(t+1) = \text{sign}\left(\sum_{j=1}^N J_{ij} S_j - \theta_i\right) \quad (3.5)$$

onde a função $\text{sign}(x)$ é definida por:

$$\text{sign}(x) = \{ +1 \text{ se } x > 0, 0 \text{ se } x = 0, -1 \text{ se } x < 0 \} \quad (3.6)$$

Esta não-linearidade do modelo é essencial para a origem das propriedades emergentes.

No próximo passo Hopfield propõe uma das mais importantes técnicas para o estudo das redes neurais. **Se a matriz de conexão for simétrica, isto é, $J_{ij} = J_{ji}$, a quantidade**

$$E = -\frac{1}{2} \sum_{i \neq j}^N J_{ij} S_i S_j \quad (3.7)$$

análoga ao conceito de energia na Física, decresce ou permanece constante durante cada iteração do sistema até atingir um mínimo (talvez local). Logo, a dinâmica da rede, para atingir um padrão ensinado é dada por uma regra de minimizar a energia.

Teorema (50)

A dinâmica do modelo de Hopfield minimiza a seguinte quantidade, denominada de energia do modelo de Hopfield quando a matriz de conexão for simétrica

$$E = -\frac{1}{2} \sum_{i \neq j}^N J_{ij} S_i S_j$$

Prova

Como a cada incremento T no tempo t apenas um neurônio é testado, vamos supor que esse neurônio seja S_i , então a variação de energia é dada por:

$$\Delta E = \left[\sum_{j \neq i} J_{ij} S_i(t+1) S_j(t) - \sum_{j \neq i} J_{ij} S_i(t) S_j(t) \right]$$

$$\Delta E = -[S_i(t+1) - S_i(t)] \left[\sum_{j \neq i} J_{ij} S_j(t) \right]$$

Usando a equação da dinâmica

$$S_i(t+1) = \text{sign} \left(\sum_{j=1}^N J_{ij} S_j - \theta_i \right)$$

tomando $S_i(t+1) = S_i(t)$ quando o argumento da função for nulo, e ainda a identidade $|x| = x \text{sign}(x)$ e

$$h_1(t) = \sum_{j \neq 1} J_{1j} S_j(t)$$

obtemos:

$$\Delta E = -[\text{sign}(\sum_{j \neq 1} J_{1j} S_j(t)) - S_1(t)] [\sum_{j \neq 1} J_{1j} S_j(t)]$$

$$\Delta E = -[\sum_{j \neq 1} J_{1j} S_j(t) \text{sign}(\sum_{j \neq 1} J_{1j} S_j(t)) - S_1(t) \sum_{j \neq 1} J_{1j} S_j(t)]$$

logo:

$$\Delta E = [|h_1(t)| - h_1(t) S_1(t)]$$

$$\Delta E = -|h_1(t)| + h_1(t) S_1(t)$$

e como

$$h_1(t) S_1(t) \leq |h_1(t)|$$

concluimos que $\Delta E \leq 0$

E Hopfield conclui **"Este caso é isomórfico ao modelo de Ising, no qual J_{ij} faz o papel da matriz de acoplamento ou de troca, com um campo externo agindo sobre cada spin. Quando J_{ij} é simétrica e possui um caráter aleatório o modelo é conhecido como vidro de spins (spin glass), dos quais sabemos que possuem muitos estados estáveis (locais)".**

Tal analogia propiciou que todo um conjunto de resultados obtidos para os vidros de spins fossem transferidos para o estudo das redes neurais, possibilitando que a Física pudesse dar sua contribuição para a área.

Ainda, em seu trabalho, Hopfield utilizando-se de simulações em computadores concluiu que o número de memórias armazenadas na rede é cerca de 15% do número de neurônios da rede.

3.2 Padrões aprendidos como atratores da rede.

No modelo de Hopfield procura-se armazenar um conjunto de p padrões $\{\xi_i^\mu\}$ $\mu=1,2,\dots,p$ onde ξ_i^μ é o i -ésimo componente do μ -ésimo padrão, podendo assumir valores ± 1 , numa rede de N neurônios.

A dinâmica da rede é dada por:

$$S_i(t+1) = \text{sign}\left(\sum_{j=1}^N J_{ij}S_j(t) - \theta_i\right) \quad (3.8)$$

Escolhendo o limiar dos neurônios como nulos:

$$S_i(t+1) = \text{sign}\left[\sum_{j=1}^N J_{ij}S_j(t)\right] \quad (3.9)$$

Vamos ainda, em vez de especificar um conjunto particular de padrões a serem memorizados, examinar o problema de forma mais genérica tomando um conjunto de padrões aleatórios dados por distribuição na qual os padrões são constituídos de bits independentes ξ_i com valores ± 1 com igual probabilidade.

Vamos agora verificar se os padrões a serem armazenados são estáveis e se pequenos desvios são corrigidos à medida que a rede desenvolve-se temporalmente.

No caso mais simples, supomos que a rede possua um único padrão memorizado. A condição para que esse padrão seja estável é que:

$$\text{sign}\left(\sum_{j=1}^N J_{ij}\xi_j\right) = \xi_i \quad \text{para todo } i \quad (3.10)$$

pois então pela dinâmica obteremos $S_i = \xi_i$.

Se tomarmos

$$J_{ij} \propto \xi_i \xi_j \quad (3.11)$$

e como $\xi_j^2 = 1$ vemos que a dinâmica será satisfeita. Tomando por conveniência a constante de proporcionalidade igual a $1/N$ obtemos:

$$J_{ij} = \frac{1}{N} \xi_i \xi_j \quad (3.12)$$

Se agora entrarmos com um padrão em que a maioria de seus bits é igual a do padrão ensinado, então a soma:

$$h_i = \sum_{j=1}^N J_{ij} S_j \quad (3.13)$$

produzirá um valor tal que $\text{sign}(h_i) = \xi_i^v$. Isto significa que uma configuração inicial próxima (no sentido de que a distância de Hamming seja pequena) a ξ^v deverá ser conduzida pela dinâmica para o padrão aprendido ξ^v que será um atrator da rede.

Ao ensinarmos o padrão ξ^v a rede aprende também o padrão $-\xi^v$, que é o reverso do padrão inicial. Todos os padrões com mais da metade de seus bits iguais ao padrão ensinado serão conduzidos a ξ^v , caso contrário para $-\xi^v$. O espaço das configurações do sistema é então simetricamente dividido em duas bases de atração.

Se ensinarmos mais de um padrão, a matriz de conexão é dada pela superposição de termos análogos para um único padrão, ou seja:

$$J_{ij} = \frac{1}{N} \sum_{\mu=1}^p \xi_i^\mu \xi_j^\mu \quad (3.14)$$

Essa regra é uma generalização da regra de Hebb na qual as mudanças sinápticas são proporcionais às correlações entre as atividades dos neurônios pré e pós

sinápticos. Entretanto, a regra acima prevê mudança nos pesos sinápticos mesmo quando não existe nenhuma atividade sináptica ($\xi_i^\mu = \xi_j^\mu = -1$). Este argumento não é fisiologicamente razoável e não possui evidência experimental. É possível modificar a regra generalizada de Hebb a fim de corrigir tais fatos (51).

Vamos agora examinar a estabilidade de um padrão particular ξ^v . Ao generalizarmos a condição de estabilidade obtemos:

$$\text{sign}(h_i^v) = \xi_i^v \quad \text{para todo } i \quad (3.15)$$

onde

$$h_i^v = \sum_{j=1}^N J_{ij} \xi_j^v = \frac{1}{N} \sum_{j=1}^N \sum_{\mu=1}^P \xi_i^\mu \xi_j^\mu \xi_j^v \quad (3.16)$$

Separando a soma em μ para o termo especial v obtemos:

$$h_i^v = \xi_i^v + \frac{1}{N} \sum_{j=1}^N \sum_{\mu \neq v}^P \xi_i^\mu \xi_j^\mu \xi_j^v \quad (3.17)$$

Se o segundo termo do lado direito da equação (3.17) é nulo ou menor que 1, concluímos que o padrão v é estável. Se assumirmos que esse termo é muito pequeno, para todo i e v , então os padrões armazenados são estáveis, isto é, se iniciarmos o sistema com um destes padrões, ele permanecerá nesse mesmo padrão de entrada. Se iniciarmos com um padrão que possui uma pequena fração de bits diferente de um dos padrões armazenados, então os bits corretos sobrepor-se-ão aos errados na soma

$$h_i = \sum_{j=1}^N J_{ij} S_j \quad (3.18)$$

e um padrão próximo a ξ^v , será conduzido através da dinâmica da rede para ξ^v . Este fato mostra que os **padrões aprendidos serão atratores da rede** e o sistema agirá como uma memória de conteúdo endereçável ou memória associativa.

Uma das contribuições mais importantes de Hopfield, foi a introdução da função energia da rede neural.

Se considerarmos

$$H = -\frac{1}{2} \sum_{i \neq j}^N J_{ij} S_i S_j \quad J_{ii} = 0 \quad (3.19)$$

então, H é função da configuração $\{S_i\}$ da rede, onde $\{S_i\}$ significa o conjunto de todos os S_i 's. A principal característica desta função é que ela sempre decresce (ou permanece constante) à medida que o sistema evolui temporalmente.

Assim os atratores são mínimos locais na superfície de energia no espaço das configurações da rede.

A dinâmica pode então ser pensada como similar ao movimento de uma partícula numa superfície de energia sob a influência da gravidade e do atrito. Então, de um ponto inicial, a partícula atingirá o repouso num ponto de mínimo local desta superfície (atrator). A base de atração corresponde a um vale desta superfície em torno do atrator. Iniciando o sistema num particular vale, a dinâmica conduzirá a partícula para o ponto de menor energia deste vale.

O conceito da função energia tem um grande número de aplicações. Em muitos campos existe uma função do estado do sistema que sempre decresce durante a evolução do sistema, ou que deve ser minimizada para obter um estado de equilíbrio.

O nome mais geral desta função, que tem origem na teoria dos sistemas dinâmicos é, função de Lyapunov (52). Outros nomes são: Hamiltoniano na Mecânica Estatística, função de custo na teoria da otimização, etc....

Para as redes neurais a função energia existe se as conexões são simétricas, isto é, $J_{ij} = J_{ji}$, que é uma suposição não condizente com a realidade. Entretanto, este fato possibilita o estudo das redes neurais por métodos matemáticos e físicos, colaborando enfim, para o estudo de redes mais complexas e mais próximas do modelo real. Em casos de conexões assimétricas, a função de Lyapunov precisa ser encontrada, o que não é uma tarefa fácil. Sabe-se entretanto, que se o sistema é estável, haverá uma função de Lyapunov que decresce com a evolução dinâmica do sistema.

3.2 Analogia spin-neurônio.

Existe uma grande analogia entre a rede de Hopfield e os vidros de spins (Apêndice A). No caso dos vidros de spins, cada spin é influenciado pelo campo magnético h . Este campo, consiste de um campo externo h_i^{ext} aplicado e de um campo interno produzido pelos outros spins.

Assim, o campo total é:

$$h_i = \sum_{j=1}^N J_{ij} S_j + h_i^{ext} \quad (3.20)$$

Num vidro de spin, as conexões são simétricas, ou seja, $J_{ij} = J_{ji}$ e estes podem assumir valores positivos, negativos ou nulos.

O campo magnético no spin S_i controla a sua dinâmica. Para baixas temperaturas, o spin tende a alinhar-se paralelamente ao campo local h_i , portanto:

$$S_i = \text{sign}(h_i) \quad (3.21)$$

Este processo, é assíncrono e randômico. A energia correspondente é dada por:

$$H = -\frac{1}{2} \sum_{i \neq j}^N J_{ij} S_i S_j - h_i^{ext} \sum_{i=1}^N S_i \quad (3.22)$$

Então, a analogia com a rede de Hopfield está completa

A conexão entre os neurônios corresponde à interação de troca entre os spins, a atividade sináptica corresponde ao campo local no spin e o campo externo corresponde ao limiar do neurônio. A função energia dos neurônios é igual à dos spins com $h_i^{ext} = 0$. A dinâmica dos spins de alinharem-se com seus campos locais, equivale a regra de MacCulloch-Pitts.

Se a temperatura, no caso dos spins, é alta, a flutuação térmica tende a destruir o alinhamento dos spins com seus campos locais. O equivalente à temperatura para o caso das redes neurais será visto adiante neste mesmo item.

O modo convencional de descrever o efeito da flutuação térmica no caso dos spins é através da dinâmica de Glauber (53) na qual a regra determinística $S_i = \text{sign}(h_i)$ é trocada por uma regra estocástica definida por:

$$S_i = \begin{cases} +1 & \text{com probabilidade } g(h_i) \\ -1 & \text{com probabilidade } 1 - g(h_i) \end{cases} \quad (3.23)$$

A função $g(h)$ depende da temperatura e é escolhida como:

$$g(h) = f_\beta(h) = \frac{1}{1 + \exp(-2\beta h)} \quad (3.24)$$

onde o parâmetro β é definido por:

$$\beta = \frac{1}{k_B T} \quad (3.25)$$

(k_B é a constante de Boltzmann ($k_B = 1,38 \cdot 10^{-16}$ erg/K))

Assim

$$\text{Prob}(S_i = \pm 1) = f_\beta(\pm h_i) = \frac{1}{1 + \exp(\mp 2\beta h_i)} \quad (3.26)$$

A temperatura controla o declive da função $f_\beta(h)$ próximo de $h = 0$ (FIGURA 10). Assim, para altas temperaturas (pequeno β) $f_\beta(h)$ muda suavemente de 0 para 1 a medida que h muda de $-\infty$ para ∞ . Já para baixas temperaturas (grandes valores de β) a mudança é mais abrupta. No limite $T \rightarrow 0$ ($\beta \rightarrow \infty$) $f_\beta(h)$ reduz-se à função degrau e recuperamos novamente a regra determinística $S_i = \text{sign}(h_i)$. No outro limite $T \rightarrow \infty$ ($\beta \rightarrow 0$) a decisão sobre o próximo valor de S_i é completamente aleatória.

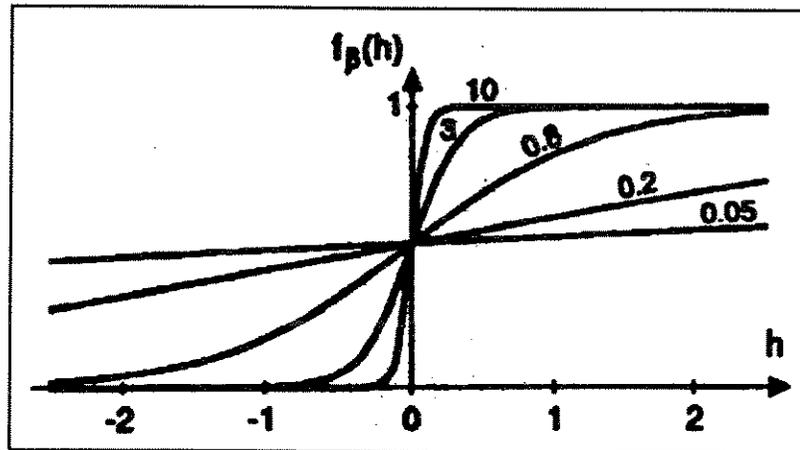


Figura 10 - Função $f(h)$ para alguns valores de β

Vamos aplicar a dinâmica de Glauber ao caso mais simples que é o de um único spin sujeito a um campo magnético externo. Então, ao calcularmos o valor médio de S (magnetização média) obtemos:

$$\langle S \rangle = (+1) \text{Prob}(S = +1) + (-1) \text{Prob}(-1) \quad (3.27)$$

$$\langle S \rangle = \frac{1}{1 + \exp(-2\beta h)} - \frac{1}{1 + \exp(2\beta h)} = \text{tgh}(\beta h) \quad (3.28)$$

A função $\text{tgh}(x)$ possui a mesma forma da sigmóide ($\text{tgh}(\beta h) = 2f_\beta(h) - 1$), mas seus limites são ± 1 .

Já num problema de muitos spins, a evolução de um spin S_i depende do campo local h_i que depende das variáveis S_j que por sua vez também sofrem o efeito da flutuação térmica. A fim de resolver este problema, faz-se uma aproximação conhecida como Teoria do Campo Médio (TCM) (54) na qual o campo local h_i é substituído por um valor médio dado por:

$$\langle h_i \rangle = \sum_{j=1}^N J_{ij} \langle S_j \rangle + h^{\text{ext}} \quad (3.29)$$

Então:

$$\langle S_i \rangle = \text{tgh}(\beta \langle h_i \rangle) = \text{tgh}(\beta \sum_{j=1}^N J_{ij} \langle S_j \rangle + \beta h^{\text{ext}}) \quad (3.30)$$

As equações acima formam um conjunto de N equações não-lineares com N variáveis ($\langle S_i \rangle$), mas que não envolve variáveis estocásticas.

O ponto principal da TCM é o de trocarmos o campo local por um campo médio tal que as flutuações dos outros spins não são consideradas, nem mesmo em resposta à mudança de estado de um determinado spin.

A aproximação da TCM torna-se exata no limite de interações de alcance infinito, na qual cada spin interage com todos os outros. Mesmo para interações de curto alcance em que $J_{ij} \approx 0$ se os spins i e j estão separados por uma distância pequena, em termos dos parâmetros da rede, a TCM nos dá uma boa descrição do fenômeno.

Vamos tomar como exemplo, uma substância ferromagnética na qual todos os J_{ij} são positivos. Assim, as interações tendem a alinhar os spins uns com os outros enquanto a flutuação térmica tende a desalinhá-los. Na ausência de um campo magnético externo, existirá uma temperatura crítica T_c , acima da qual a flutuação térmica destrói toda ordem e obtemos $\langle S \rangle = 0$ enquanto que abaixo de T_c temos $\langle S \rangle \neq 0$. O sistema exibirá uma transição de fase na temperatura T_c . O ferro é um material cuja temperatura crítica é $770^\circ C$. Acima desta temperatura o ferro perde sua magnetização. No modelo mais simples de ferromagnetismo tomamos:

$$J_{ij} = \frac{J}{N} \quad \text{para todo } i,j \quad (3.31)$$

onde J é constante e o fator $1/N$ é necessário para que a energia do sistema permaneça uma propriedade extensiva do sistema.

O modelo representado pela equação (3.31), equivale ao de Hopfield no qual armazenamos um único padrão $\xi_i = 1$ para todo i. Então, pela sua dinâmica o modelo dirigir-se-á para o padrão ensinado, ou ao seu reverso dependendo do número de bits iniciais iguais entre o padrão de entrada e o armazenado.

Se ensinarmos à rede um único padrão ξ_i , com valores ± 1 para seus componentes podemos definir novas variáveis (55) por:

$$S_i' = \xi_i S_i \quad (3.32)$$

e observamos que o modelo citado comporta-se como um ferromagneto em relação ao padrão ensinado, isto é, através de sua dinâmica ele dirigir-se-á para o padrão ξ_i ou ao seu reverso.

Destacamos o fato de que, a dinâmica do processo de reconhecimento numa rede neural equivale ao realinhamento dos spins no material ferromagnético ao campo magnético criado pelos seus vizinhos, até que todos apontem na mesma direção.

Quando generalizamos para o caso do aprendizado de vários padrões, podemos imaginar a rede neural como um material ferromagnético que possui um número de padrões de magnetização que são localmente estáveis (vidros de spins). Portanto, a relaxação para qualquer uma das memórias da rede equivale no vidro de spin para um estado magnético estável.

Logo, para um material ferromagnético, que é caracterizado por possuir uma magnetização uniforme, ou seja,

$$\langle S_i \rangle = \langle S \rangle \quad (3.33)$$

obtemos das equações (3.30) e (3.33)

$$\langle S \rangle = \tanh(\beta J \langle S \rangle) \quad (3.34)$$

A equação acima pode ser resolvida graficamente e verificamos que quando $\beta J \leq 1$ a única solução é $\langle S \rangle = 0$, FIGURA 11 (a), denominada de fase paramagnética, enquanto que para $\beta J > 1$ existem 3 soluções, FIGURA 11 (b). Logo

$$T_c = \frac{J}{k_B} \quad (3.35)$$

é a temperatura crítica. Das 3 soluções aquela com magnetização nula é instável enquanto as outras duas diferem apenas no sinal.

Observamos que à medida que $T \rightarrow 0$ ($\beta \rightarrow \infty$) recuperamos o modelo determinístico com a função tangente hiperbólica sendo substituída pela função sinal.

Introduziremos agora o conceito análogo ao da temperatura para o caso das **redes neurais**. Tal analogia permitirá fazer uso da TCM para as redes neurais. Tomando o limite $T \rightarrow 0$ deveremos também, recuperar o modelo de Hopfield.

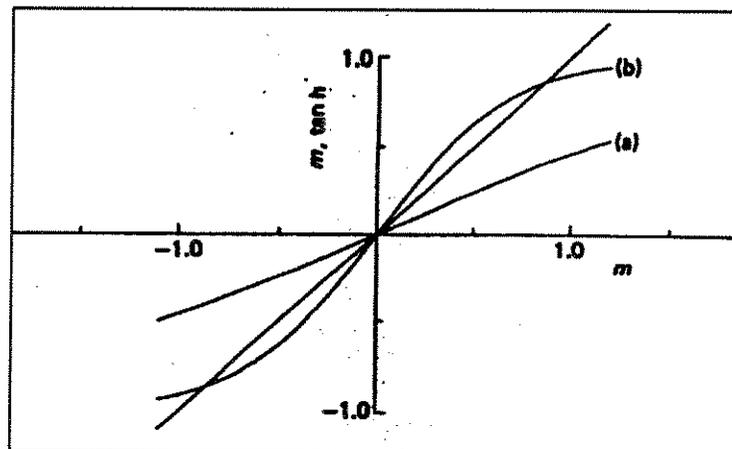


Figura 11 - Solução Gráfica para a Magnetização - (Ref. (70))

Vamos assumir que os neurônios da rede comportam-se de forma estocástica como no caso dos spins, e (56,57) que a probabilidade de que S_i escolhido aleatoriamente dispare ou não, é dada por:

$$\text{Prob}(S_i = \pm 1) = f_{\beta}(\pm h_i) = \frac{1}{1 + \exp(\mp 2\beta h_i)} \quad (3.36)$$

Então, obteremos de forma análoga aos vidros de spins que para um neurônio estocástico, seu valor médio é expresso por:

$$\langle S_i \rangle = \text{tgh}(\beta h_i) \quad (3.37)$$

Esta escolha torna possível o uso da Mecânica Estatística no estudo das redes neurais, e ainda nos permitirá o cálculo dos valores médios das grandezas envolvidas nos parâmetros de ordem das redes neurais no próximo capítulo.

O significado deste comportamento estocástico para os neurônios, é devido ao fato destes sofrerem flutuações randômicas no lançamento dos neurotransmissores, ao atraso dos sinais nas sinapses etc....

Estes efeitos provocam um ruído o qual podemos descrever analogamente a uma flutuação térmica. **Evidentemente, β , no caso das redes neurais não é a temperatura física do sistema, é apenas um parâmetro controlando o nível de ruído e é definido por:**

$$\beta = \frac{1}{T} \quad (3.38)$$

Então, para a temperatura $T \rightarrow 0$ a sigmóide reduz-se à função degrau e recuperamos o modelo de neurônio determinístico do tipo MacCulloch-Pitts.

O uso do neurônio com comportamento estocástico permite a rede sair de estados espúrios, que são mínimos locais da função energia. Geralmente, estes estados são menos estáveis que os padrões armazenados na rede.

A rede evoluirá de forma diferente cada vez que é usada. As propriedades significativas calculadas, são médias sobre todas as possíveis evoluções, ponderadas pelas probabilidades de cada história particular. Mas, para que possamos nos utilizar da Mecânica Estatística é necessário que a rede atinja um estado de equilíbrio, ou seja, que os valores médios de grandezas tais como S_i , isto é $\langle S_i \rangle$ tornem-se valores independentes do tempo. Como já vimos, redes com função energia com conexões simétricas atingem um estado de equilíbrio.

Assim, devemos trabalhar não com a estabilidade absoluta de uma dada configuração $\{S_i\}$ mas sim, com estados de equilíbrio estáveis $\{\langle S_i \rangle\}$ nos quais os valores médios não mudam com o tempo.

CAPÍTULO 4.

MÉTODO DOS CUMULANTES

O propósito deste capítulo, é obtermos de forma simples os parâmetros de ordem de redes neurais sem levarmos em conta os métodos da Mecânica Estatística. Para isto, faremos a expansão cumulante até segunda ordem.

4.1 Método dos cumulantes

Iniciaremos por rever uns poucos conceitos de Estatística que utilizaremos nesse item. Para um estudo aprofundado sugerimos as referências (58,59,60-62).

Seja X uma variável aleatória e $P(x)$ sua função densidade de probabilidade, isto é, $P(x)dx$ é a probabilidade da variável X tomar valores no intervalo $(x, x+dx)$, logo:

$$P(x)dx \equiv \text{Prob}\{X \in [x, x+dx]\} \quad (4.1)$$

Então da definição acima, e da lei da adição das probabilidades, obtemos:

$$\int_a^b P(x)dx = \text{Prob}\{X \in [a,b]\} \quad (4.2)$$

A função densidade de probabilidade deve ser não negativa e satisfazer

$$\int_{-\infty}^{\infty} P(x)dx = 1 \quad (4.3)$$

A função de distribuição de probabilidade é definida por:

$$F(x) \equiv \text{Prob}\{X < x\} \quad (4.4)$$

Da equação acima obtemos:

$$F(x) = \int_{-\infty}^x P(\xi)d\xi \quad (4.5)$$

Seja agora $f(x)$ uma função real. Então, o valor médio de $f(x)$ em relação a variável X é:

$$\langle f(x) \rangle = \int_{-\infty}^{\infty} f(x)P(x)dx \quad (4.6)$$

O valor médio da variável X é definido por:

$$\mu \equiv \langle X \rangle = \int_{-\infty}^{\infty} xP(x)dx \quad (4.7)$$

e sua variância:

$$\sigma^2 \equiv \langle (X - \mu)^2 \rangle = \int_{-\infty}^{\infty} (x - \mu)^2 P(x)dx = \langle X^2 \rangle - \mu^2 \quad (4.8)$$

O enésimo momento da variável X é expresso por:

$$\langle X^n \rangle = \int_{-\infty}^{\infty} x^n P(x) dx \quad (4.9)$$

O primeiro momento é o valor médio de X.

A transformada de Fourier (63) da função densidade de probabilidade é denominada de função característica da variável X e é dada por:

$$G_X(k) = \int_{-\infty}^{\infty} e^{ikx} P(x) dx \quad (4.10)$$

Desenvolvendo a exponencial em série de Taylor e efetuando as integrais obtemos:

$$G_X(k) = \sum_{n=0}^{\infty} \frac{(ik)^n}{n!} \langle X^n \rangle \quad (4.11)$$

A transformada inversa da equação (4.10) nos dá:

$$P(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-ikx} G_X(k) dk \quad (4.12)$$

As equações (4.11) e (4.12) mostram que podemos reconstruir a função densidade de probabilidade se conhecermos todos os momentos da variável X, e se os mesmos forem finitos.

Podemos escrever (4.11) como:

$$G_X(k) = 1 + \sum_{n=1}^{\infty} \frac{(ik)^n}{n!} \langle X^n \rangle \quad (4.13)$$

Para a função $\ln(1+z)$ temos a expansão:

$$\ln(1+z) = \frac{z}{1} - \frac{z^2}{2} + \frac{z^3}{3} - \dots \quad (4.14)$$

OBS:

Como $|e^{ikx}| \leq 1$ para todo x observamos que:

$$\int_{-\infty}^{\infty} |e^{ikx} P(x)| dx \leq \int_{-\infty}^{\infty} P(x) dx = 1$$

Então tomando o logaritmo natural de $G_X(k)$ obtemos:

$$\ln G_X(k) = \sum_{n=1}^{\infty} \frac{(ik)^n}{n!} \langle X^n \rangle - \frac{1}{2} \left[\sum_{n=1}^{\infty} \frac{(ik)^n}{n!} \langle X^n \rangle \right]^2 + \dots \quad (4.15)$$

Se agora rearranjarmos os termos em potências de k teremos

$$\ln G_X(k) = \ln \left[1 + \sum_{n=1}^{\infty} \frac{(ik)^n}{n!} \langle X^n \rangle \right] = \sum_{n=1}^{\infty} \frac{(ik)^n}{n!} C_n(X) \quad (4.16)$$

ou ainda:

$$G_X(k) = 1 + \sum_{n=1}^{\infty} \frac{(ik)^n}{n!} \langle X^n \rangle = \exp \left[\sum_{n=1}^{\infty} \frac{(ik)^n}{n!} C_n(X) \right] \quad (4.17)$$

portanto:

$$\ln G_X(k) = \ln \left(1 + (ik)\mu_1 + \frac{(ik)^2 \mu_2}{2!} + \frac{(ik)^3 \mu_3}{3!} + \dots \right)$$

$$\ln G_X(k) = (ik)\mu_1 + \frac{(ik)^2}{2!} (\mu_2 - \mu_1^2) + \frac{(ik)^3}{3!} (\mu_3 - 3\mu_1\mu_2 + 2\mu_1^3) + \dots$$

$$\ln G_X(k) = (ik)C_1(X) + \frac{(ik)^2 C_2(X)}{2!} + \frac{(ik)^3 C_3(X)}{3!} + \dots \quad (4.18)$$

Os coeficientes $C_n(X)$ foram introduzidos por Thiele (64) e são denominados de cumulantes ou semi-invariantes da função densidade de probabilidade $P(x)$.

Os primeiros quatro cumulantes são expressos por:

$$\begin{aligned} C_1(X) &= \langle X \rangle \\ C_2(X) &= \langle X^2 \rangle - \langle X \rangle^2 \\ C_3(X) &= \langle X^3 \rangle - 3\langle X \rangle \langle X^2 \rangle + 2\langle X \rangle^3 \\ C_4(X) &= \langle X^4 \rangle - 3\langle X^2 \rangle^2 - 4\langle X \rangle \langle X^3 \rangle + 12\langle X \rangle^2 \langle X^2 \rangle - 6\langle X \rangle^4 \end{aligned} \quad (4.19)$$

Os três primeiros cumulantes são interpretados como:

O primeiro é o valor médio da variável X . O segundo a variância da variável X que nos dá uma medida da dispersão da variável em torno de seu valor médio. O terceiro quantifica a assimetria da função densidade de probabilidade.

Seja a função densidade de probabilidade normal com valor médio da variável X dado por μ e variância σ^2 , então:

$$P(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \quad (4.20)$$

Verificamos que:

$$\langle x \rangle = \int_{-\infty}^{\infty} x P(x) dx = \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^{\infty} x e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} (\mu + \sigma x) e^{-\frac{x^2}{2}} dx = \mu \quad (4.21)$$

e ainda

$$\langle X - \mu \rangle^2 = \int_{-\infty}^{\infty} (x - \mu)^2 P(x) dx = \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^{\infty} (x - \mu)^2 e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx = \frac{\sigma^2}{\sqrt{2\pi}} \int_{-\infty}^{\infty} x^2 e^{-\frac{x^2}{2}} dx = \sigma^2 \quad (4.22)$$

Dada a função característica, podemos invertê-la para encontrarmos a função densidade da variável X.

A função delta de Dirac é definida por:

$$\delta(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{ikx} dk \quad (4.23)$$

então:

$$\begin{aligned} \frac{1}{2\pi} \int_{-\infty}^{\infty} G_X(k) e^{-ikx} dk &= \frac{1}{2\pi} \int_{k=-\infty}^{\infty} \left[\int_{y=-\infty}^{\infty} e^{iky} P(y) dy \right] e^{-ikx} dk \\ &= \int_{y=-\infty}^{\infty} \left[\frac{1}{2\pi} \int_{k=-\infty}^{\infty} e^{ik(y-x)} dk \right] P(y) dy \\ &= \int_{-\infty}^{\infty} \delta(y-x) P(y) dy = P(x) \end{aligned} \quad (4.24)$$

Quando a função densidade é uma função par:

$$G_X(k) = \int_{-\infty}^{\infty} e^{ikx} P(x) dx = \int_{-\infty}^{\infty} \cos(kx) P(x) dx \quad (4.25)$$

assim a função característica será real.

Podemos obter da função característica os momentos da variável X por diferenciação. Então:

$$\frac{d}{dk} [G_X(k)] = i \int_{-\infty}^{\infty} x e^{ikx} P(x) dx \quad (4.26)$$

$$\frac{d}{dk} [G_X(k)]_{k=0} = i \int_{-\infty}^{\infty} x P(x) dx = i\mu \quad (4.27)$$

$$\frac{d^2}{dk^2} [G_X(k)] = i^2 \int_{-\infty}^{\infty} x^2 e^{ikx} P(x) dx \quad (4.28)$$

$$\frac{d^2}{dk^2} [G_X(k)]_{k=0} = i^2 \int_{-\infty}^{\infty} x^2 P(x) dx = i^2 \langle X^2 \rangle \quad (4.29)$$

logo:

$$\frac{d^n}{dk^n} [G_X(k)]_{k=0} = i^n \langle X^n \rangle \quad (4.30)$$

Existem fórmulas para o cálculo dos cumulantes de ordem n (65-69)

$$C_n(X) = -n! \sum_{\{p_j\}} \left\{ \left[\sum_j (p_j - 1)! \right] \prod_j \frac{(-1)^{p_j} \langle X^j \rangle^{p_j}}{p_j! (j!)^{p_j}} \right\} \quad (4.31)$$

onde a primeira somatória é sobre todos os conjuntos de inteiros não negativos tais que:

$$\sum_{j=1}^n j p_j = n \quad (4.32)$$

Se a função densidade for gaussiana então:

$$G_X(k) = \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^{\infty} e^{ikx} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx \quad (4.33)$$

definindo

$$r = \frac{x-\mu}{\sigma} \quad (4.34)$$

obtemos:

$$G_X(k) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{1}{2}r^2 + ik\sigma r} dr = \frac{1}{\sqrt{2\pi}} e^{ik\mu} \int_{-\infty}^{\infty} e^{-\frac{1}{2}(r-ik\sigma)^2 + k^2\sigma^2} dr \quad (4.35)$$

ou ainda,

$$G_X(k) = \frac{1}{\sqrt{2\pi}} e^{ik\mu - \frac{1}{2}k^2\sigma^2} \int_{-\infty}^{\infty} e^{-\frac{1}{2}(r-ik\sigma)^2} dr \quad (4.36)$$

e fazendo $z = r - ik\sigma$:

$$G_X(k) = \frac{1}{\sqrt{2\pi}} e^{ik\mu - \frac{1}{2}k^2\sigma^2} \int_{-\infty}^{\infty} e^{-\frac{z^2}{2}} dz = e^{ik\mu - \frac{1}{2}k^2\sigma^2} \quad (4.37)$$

Obtemos então os cumulantes de uma função de densidade normal como:

$$\begin{aligned} C_1(X) &= \mu \\ C_2(X) &= \sigma^2 \\ C_3(X) &= 0 \\ C_4(X) &= 0 \end{aligned} \quad (4.38)$$

Logo, para a função densidade normal, só obtemos os dois primeiros cumulantes.

Média de uma função $f(x)$

Para uma função $f(x)$, seu valor médio é calculado por:

$$\langle f(x) \rangle = \int_{-\infty}^{\infty} f(x) P(x) dx = \int_{-\infty}^{\infty} f(x) dx \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-ikx} G_X(k) dk \quad (4.39)$$

Se agora desenvolvermos a função característica até seu segundo cumulante, teremos:

$$\langle f(x) \rangle = \int_{-\infty}^{\infty} f(x) dx \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-ikx} e^{(ik)C_1 + \frac{(ik)^2}{2}C_2} dk \quad (4.40)$$

ou após uma transformação direta

$$\langle f(x) \rangle = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{z^2}{2}} f(\sqrt{C_2}z + C_1) dz \quad (4.41)$$

A equação acima será a fórmula básica para o cálculo dos valores médios das grandezas envolvidas nos parâmetros de ordem das redes neurais, pois através da equação podemos calcular o valor médio de uma função em termos de seu primeiro e segundo cumulante. Observamos que para o caso da função densidade normal, a fórmula acima é exata.

4.2 - Aplicação do método ao modelo de Hopfield

Vamos considerar uma rede neural do tipo Hopfield com N neurônios estocásticos com a regra de aprendizado generalizada de Hebb, na qual p padrões $\{\xi^\mu\}$ $\mu=1,2,\dots,p$ foram ensinados, com capacidade de armazenamento $\alpha = p/N$, com α finito. Quando o número de padrões ensinados à rede tende ao infinito, existe a possibilidade de que a sobreposição média entre um estado inicial qualquer da rede e os padrões ensinados se anulem. A fim de excluir este caso, assumimos que existe um número finito de estados que possuem uma sobreposição diferente de zero denominados de **padrões condensados**. Vamos considerar que os padrões de $\mu=1,\dots,s$ com $1 < s < p$ irão condensar-se enquanto o restante $\mu=s+1,\dots,p$ não.

O campo total no neurônio i devido à mútua influência entre os neurônios é:

$$h_i = \frac{1}{N} \sum_{j \neq i}^N J_{ij} \langle S_j \rangle \quad (4.42)$$

e este campo produz um valor para o estado do neurônio i dado por:

$$\langle S_i \rangle = \tanh(\beta h_i) \quad (4.43)$$

Introduzindo os valores de J_{ij} dados pela regra de Hebb na equação (4.42), e definindo a sobreposição média m^μ , entre um padrão ensinado e um estado inicial qualquer da rede por:

$$m^\mu = \frac{1}{N} \sum_{j=1}^N \xi_j^\mu \langle S_j \rangle \quad (4.44)$$

obtemos:

$$h_i = \sum_{\mu=1}^p h_i^\mu - \alpha \langle S_i \rangle \quad (4.45)$$

onde:

$$h_i^\mu = \xi_i^\mu m^\mu \quad (4.46)$$

No limite termodinâmico ($N \rightarrow \infty$) podemos desprezar o último termo da equação (4.45) acima.

Vimos que a média de uma função $f(x)$ de uma distribuição gaussiana é dada por:

$$\langle f(x) \rangle = \int_{-\infty}^{\infty} \frac{dz}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} f(C_1 + \sqrt{C_2} z) \quad (4.47)$$

Na equação (3.37) se identificarmos $f(x)$ com $\langle S_i \rangle$ teremos:

$$f(\beta h_i) = \text{tgh}(\beta h_i) \quad (4.48)$$

Então

$$f(C_1 + \sqrt{C_2} z) = \text{tgh}[\beta (\langle h_i \rangle + \sqrt{\langle h_i^2 \rangle - \langle h_i \rangle^2} z)] \quad (4.49)$$

Vamos agora assumir que o campo num neurônio i pode ser expresso como um valor médio mais uma pequena variação, da seguinte forma,:

$$h_i = \langle h_i \rangle + \delta h_i \quad (4.50)$$

Verificamos que $\langle (\delta h_i)^2 \rangle$ pode ser expresso por:

$$\langle (\delta h_i)^2 \rangle = \langle h_i^2 \rangle - \langle h_i \rangle^2 \quad (4.51)$$

pois,

$$\langle (\delta h_i)^2 \rangle = \langle (h_i - \langle h_i \rangle)^2 \rangle = \langle h_i^2 - 2h_i \langle h_i \rangle + \langle h_i \rangle^2 \rangle \quad (4.52)$$

Podemos ainda decompor o campo h_i em duas partes, a parte dos padrões que condensam e a parte dos padrões com $\mu > s$, então:

$$h_i = \langle \sum_{\mu=1}^s h_i^\mu \rangle + \sum_{\mu>s} h_i^\mu \quad (4.53)$$

Assim:

$$\langle h_i \rangle = \langle \sum_{\mu=1}^s \xi_i^\mu m^\mu \rangle \quad (4.54)$$

e obtemos para

$$\langle (\delta h_i)^2 \rangle = \langle \left(\sum_{\mu>s} \xi_i^\mu m^\mu \right)^2 \rangle = \langle \left(\sum_{\mu>s} m^\mu \right)^2 \rangle \quad (4.55)$$

Das equações acima concluímos:

$$f(C_1 + \sqrt{C_2} z) = \text{tgh} \left[\beta \left(\langle \sum_{\mu=1}^s \xi_i^\mu m^\mu \rangle + \sqrt{\langle \sum_{\mu>s} (m^\mu)^2 \rangle} z \right) \right] \quad (4.56)$$

Vamos agora introduzir uma variável auxiliar r para descrever o ruído devido aos padrões que não se condensam, assim, se a rede está num estado com sobreposição diferente de zero com apenas alguns padrões ensinados, a soma das sobreposições

aleatórias com todos os outros padrões, faz com que apareça uma significativa quantidade de ruído descrito pelo parâmetro r definido por Amit, (70) como $\alpha r = \langle (\delta h_i)^2 \rangle$ ou

$$\alpha r = \langle \sum_{\mu > s}^p (m^\mu)^2 \rangle \quad (4.57)$$

A partir das equações acima obtemos:

$$\langle S_i \rangle = \int_{-\infty}^{\infty} \frac{dz}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} \operatorname{tgh}[\beta (\bar{m} \cdot \bar{\xi}_i + \sqrt{\alpha} rz)] \quad (4.58)$$

$$m^\mu = \frac{1}{N} \sum_{i=1}^N \xi_i^\mu \int_{-\infty}^{\infty} \frac{dz}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} \operatorname{tgh}[\beta (\bar{m} \cdot \bar{\xi}_i + \sqrt{\alpha} rz)] \quad (4.59)$$

que podemos expressá-la por:

$$m^\mu = \langle \langle \xi_i^\mu \operatorname{tgh}[\beta (\bar{m} \cdot \bar{\xi}_i + \sqrt{\alpha} rz)] \rangle \rangle \quad (4.60)$$

onde $\langle \langle \dots \rangle \rangle$ significa duas médias, uma gaussiana e a outra sobre a distribuição dos ξ_i^μ .

Necessitamos ainda do parâmetro q para descrever a fase de vidro de spin. Este parâmetro é denominado de parâmetro de Edwards-Anderson (71). Uma explicação mais detalhada é dada no APÊNDICE A.

Como q é definido por:

$$q = \frac{1}{N} \sum_{i=1}^N \langle S_i \rangle^2 \quad (4.61)$$

para determiná-lo temos apenas que identificar $f(x)$ com o quadrado da função tangente hiperbólica, assim:

$$\langle S_i \rangle^2 = \int_{-\infty}^{\infty} \frac{dz}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} \operatorname{tgh}^2 [\beta (\bar{m} \cdot \bar{\xi}_i + \sqrt{\alpha} rz)] \quad (4.62)$$

e então

$$q = \langle \langle \operatorname{tgh}^2 [\beta (\bar{m} \cdot \bar{\xi} + \sqrt{\alpha} rz)] \rangle \rangle \quad (4.63)$$

Na determinação do parâmetro r , utilizaremos um método similar ao de Peretto (72) e Domany e outros (73). Usando as equações (4.44), (4.45) e (4.48) obtemos:

$$m^\mu = \frac{1}{N} \sum_{i=1}^N \xi_i^\mu \operatorname{tgh} [\beta (\sum_{\mu=1}^p h_i^\mu - \alpha \langle S_i \rangle)] \quad (4.64)$$

Desenvolvendo a função tangente hiperbólica em relação a h_i^μ obtemos:

$$m^\mu (h^\mu) = m^\mu (h^\mu = 0) + h^\mu \left. \frac{\partial m^\mu}{\partial h^\mu} \right|_{h^\mu=0} + \dots$$

$$m^\mu = \frac{1}{N} \sum_{i=1}^N \xi_i^\mu \operatorname{tgh} [\beta (\sum_{\nu \neq \mu}^p h_i^\nu - \alpha \langle S_i \rangle)] + \frac{1}{N} \sum_{i=1}^N \xi_i^\mu \beta h_i^\mu \{1 - \operatorname{tgh}^2 [\beta (\sum_{\nu \neq \mu}^p h_i^\nu - \alpha \langle S_i \rangle)]\} \quad (4.65)$$

Substituindo $h_i^\mu = \xi_i^\mu m^\mu$ na equação acima obtemos:

$$m^\mu [1 - \frac{\beta}{N} \sum_{i=1}^N \{1 - \operatorname{tgh}^2 [\beta (\sum_{\mu=1}^p h_i^\mu - \alpha \langle S_i \rangle)]\}] = \frac{1}{N} \sum_{i=1}^N \xi_i^\mu \operatorname{tgh} [\beta (\sum_{\mu=1}^p h_i^\mu - \alpha \langle S_i \rangle)] \quad (4.66)$$

Elevando ao quadrado a equação acima e somando sobre todos os padrões $\mu > s$ obtemos para o parâmetro r

$$r = \frac{q}{[1 - \beta(1 - q)]^2} \quad (4.67)$$

4.3 Diagrama de fase da rede de Hopfield (70,74).

Das equações (4.60), (4.63) e (4.67) obtemos para os parâmetros de ordem:

$$m^\mu = \langle\langle \xi^\mu \operatorname{tgh} \beta [\sqrt{\alpha r z} + \bar{m} \cdot \bar{\xi}] \rangle\rangle$$

$$q = \langle\langle \operatorname{tgh}^2 \beta [\sqrt{\alpha r z} + \bar{m} \cdot \bar{\xi}] \rangle\rangle$$

$$r = \frac{q}{[1 - \beta(1 - q)]^2}$$

Observações:

Usamos a representação $\langle\langle \dots \rangle\rangle$ para indicarmos a média sobre os padrões discretos condensados ξ^μ com $v \leq s$ e a média gaussiana sobre o ruído gerado pelos outros padrões. Essa média combinada pode ser expressa por:

$$\langle\langle A \rangle\rangle = \frac{1}{2^s} \sum_{\mu=1}^s \sum_{\xi^\mu = \pm 1} \int_{-\infty}^{\infty} \frac{dz}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} A\{\xi\} \quad (4.68)$$

As equações acima para m, q, r são compostas de duas partes:

1. Uma parte ferromagnética $\bar{m} \cdot \bar{\xi}$ que é devida aos s padrões condensados.
2. A outra parte devido ao ruído provocado pela fase de vidro de spin $\sqrt{\alpha r z}$ gerada pela sobreposição aleatória dos padrões armazenados.

A solução numérica e analítica das equações acima, produz o diagrama de fase da rede de Hopfield como mostra a figura abaixo.

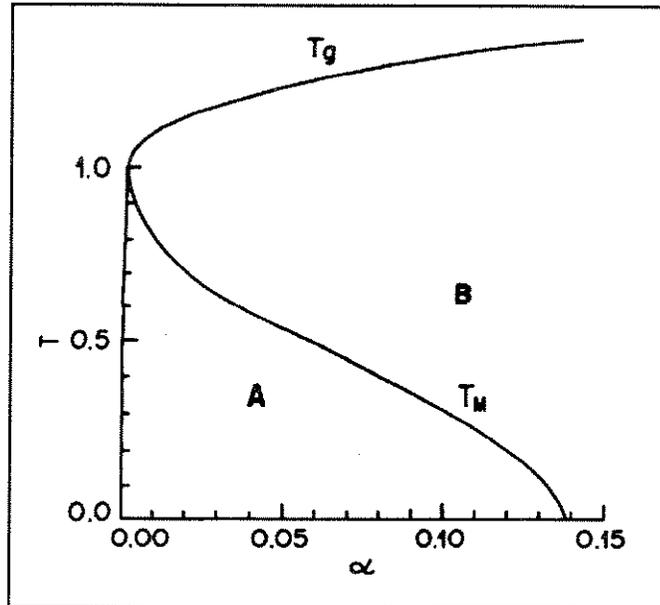


Figura 12 - Diagrama de Fase do Modelo de Hopfield

com as seguintes observações:

1. Para $\alpha < 0,138$ (fase ferromagnética) existe a recuperação dos padrões armazenados (**Região A**).

2. Para $\alpha > 0,138$ (fase de vidro de spin) não há recuperação dos padrões armazenados (**Região B**).

Análise das equações para $T \rightarrow 0$.

Consideremos o caso no qual recuperamos o padrão ξ^1 , isto é, com o vetor \bar{m} possuindo um único componente m^1 diferente de zero. Tomando ainda, o campo externo nulo ($h=0$) teremos:

$$m^1 = \int \frac{dz}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} \operatorname{tgh} \beta [m^1 + \sqrt{\alpha} r z] \quad (4.69)$$

$$q = \int \frac{dz}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} \operatorname{tgh}^2 \beta [m^1 + \sqrt{\alpha r} z] \quad (4.70)$$

$$r = \frac{q}{[1 - \beta(1 - q)]^2} \quad (4.71)$$

Então definindo:

$$C \equiv \beta(1 - q) \quad (4.72)$$

temos que:

$$C = \beta \int_{-\infty}^{\infty} \frac{dz}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} \cosh^{-2}[\beta (\sqrt{\alpha r} z + m^1)] \quad (4.73)$$

$$C = \beta \int_{-\infty}^{\infty} \frac{dz}{\sqrt{2\pi}} \exp\left\{-\frac{z^2}{2} - \ln \cosh^2[\beta (\sqrt{\alpha r} z + m^1)]\right\} \quad (4.74)$$

Uma vez que $\beta \rightarrow \infty$, desenvolvemos o argumento da exponencial próximo a

$$z = -\frac{m^1}{\sqrt{\alpha r}} \quad (4.75)$$

e obtemos:

$$C = \beta \int_{-\infty}^{\infty} \frac{dz}{\sqrt{2\pi}} \exp\left\{-\frac{z^2}{2} - \beta^2 (\sqrt{\alpha r} z + m^1)^2\right\} \quad (4.76)$$

Após integrarmos teremos:

$$C = \sqrt{\frac{2}{\pi \alpha r}} \exp\left(-\frac{m_1^2}{2 \alpha r}\right) \quad (4.77)$$

Utilizando agora a identidade:

$$\begin{aligned}
\int \frac{dz}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} \operatorname{tgh} \beta [az+b] &\xrightarrow{r \rightarrow 0} \int \frac{dz}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} \operatorname{sign} \beta [az+b] \\
&= 2 \int_{-\frac{b}{a}}^{\infty} \frac{dz}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} - 1 \\
&= \operatorname{erf}\left(\frac{b}{\sqrt{2\alpha}}\right) \tag{4.78}
\end{aligned}$$

onde a função erro é definida por:

$$\operatorname{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x \exp(-u^2) du \tag{4.79}$$

as equações dos parâmetros de ordem ficam com $m^1 = m$:

$$C \equiv \beta(1-q) = \sqrt{\frac{2}{\pi\alpha r}} \exp\left(-\frac{m^2}{2\alpha r}\right) \tag{4.80}$$

$$r = \frac{1}{(1-C)^2} \tag{4.81}$$

$$m = \operatorname{erf}\left(\frac{m}{\sqrt{2\alpha r}}\right) \tag{4.82}$$

Se fizermos:

$$y = \frac{m}{\sqrt{2\alpha r}} \tag{4.83}$$

obteremos:

$$y(\sqrt{2\alpha} + \frac{2}{\pi} e^{-y^2}) = \operatorname{erf}(y) \tag{4.84}$$

Resolvida a equação acima, obtém-se um valor crítico $\alpha_c \cong 0,138$ acima do qual m é sempre nulo.

4.4 Aplicações a outros modelos.

A- Modelo de rede neural com esquecimento (75).

Nesse modelo a interação entre os neurônios é expressa por:

$$J_{ij} = \frac{1}{N} \sum_{\mu=1}^p \Lambda\left(\frac{\mu}{N}\right) \xi_i^\mu \xi_j^\mu \quad (4.85)$$

onde $\Lambda(x)$ é a função dada por:

$$\Lambda(x) = \varepsilon \exp\left(-\frac{1}{2} \varepsilon^2 x^2\right) \quad (4.86)$$

com a condição de normalização:

$$\int_{-\infty}^{\infty} \Lambda^2(x) dx = 1 \quad (4.87)$$

Este modelo permite à rede neural, o esquecimento após o aprendizado. O padrão mais recentemente memorizado pela rede é o de $\mu=1$, e o mais antigo, é o de $\mu=p$.

Com essas considerações, a memorização e a capacidade de aprendizado da rede, dependem do parâmetro ε . Para valores de ε grandes, maiores serão os pesos dos padrões mais recentemente memorizados.

O parâmetro de ordem r é definido por:

$$r = \left\langle \left\langle \sum_{\mu=s+1}^{p-\alpha N} \Lambda^2\left(\frac{\mu}{N}\right) \langle m^\mu \rangle^2 \right\rangle \right\rangle \quad (4.88)$$

Podemos expressar o campo num dado neurônio i por:

$$h_i = \frac{1}{N} \sum_{j=1}^N \sum_{\mu=1}^p \Lambda\left(\frac{\mu}{N}\right) \xi_i^\mu \xi_j^\mu \langle S_j \rangle \quad (4.89)$$

ou usando a sobreposição média

$$h_i = \sum_{\mu=1}^p \Lambda\left(\frac{\mu}{N}\right) \xi_{s_i}^{\mu} m^{\mu} \quad (4.90)$$

Se fizermos

$$h_i^{\mu} = \Lambda\left(\frac{\mu}{N}\right) \xi_{s_i}^{\mu} m^{\mu} \quad (4.91)$$

podemos também expressar o campo por:

$$h_i = \sum_{\mu=1}^p h_i^{\mu} = \sum_{\mu=1}^s h_i^{\mu} + \sum_{\mu>s} h_i^{\mu} \quad (4.92)$$

ou como anteriormente

$$h_i = \langle h_i \rangle + \delta h_i \quad (4.93)$$

Assim, o valor do campo médio será:

$$\langle h_i \rangle = \left\langle \sum_{\mu=1}^p \Lambda\left(\frac{\mu}{N}\right) \xi_{s_i}^{\mu} m^{\mu} \right\rangle \quad (4.94)$$

e o de $\langle (\delta h_i)^2 \rangle$:

$$\langle (\delta h_i)^2 \rangle = \left\langle \left(\sum_{\mu>s} \Lambda\left(\frac{\mu}{N}\right) \xi_{s_i}^{\mu} m^{\mu} \right)^2 \right\rangle \quad (4.95)$$

ou ainda

$$\langle (\delta h_i)^2 \rangle = \left\langle \sum_{\mu>s} \Lambda^2\left(\frac{\mu}{N}\right) (m^{\mu})^2 \right\rangle \quad (4.96)$$

Definindo o parâmetro r por $r = \langle (\delta h_i)^2 \rangle$ e usando os mesmos argumentos do item 4.2 temos:

$$m^\mu = \langle \langle \xi_i^\mu \operatorname{tgh}[\beta (\sum_{\mu=1}^s \Lambda(\frac{\mu}{N}) \xi_i^\mu m^\mu + \sqrt{rz})] \rangle \rangle \quad (4.97)$$

$$q = \langle \langle \operatorname{tgh}^2[\beta (\sum_{\mu=1}^s \Lambda(\frac{\mu}{N}) \xi_i^\mu m^\mu + \sqrt{rz})] \rangle \rangle \quad (4.98)$$

O parâmetro r pode ser obtido de maneira análoga ao item anterior através de:

$$m^\mu = \frac{1}{N} \sum_{i=1}^N \xi_i^\mu [\operatorname{tgh}(\beta \sum_{\nu \neq \mu}^p \Lambda(\frac{\nu}{N}) \xi_i^\nu m^\nu) + \beta \Lambda(\frac{\mu}{N}) \xi_i^\mu m^\mu (1 - \operatorname{tgh}^2(\beta \sum_{\nu \neq \mu}^p \Lambda(\frac{\mu}{N}) \xi_i^\nu m^\nu))] \quad (4.99)$$

Então

$$\langle \sum_{\mu > s}^p \Lambda^2(\frac{\mu}{N}) (m^\mu)^2 \rangle = \sum_{\mu > s}^p \frac{q \Lambda^2(\mu/N)}{[1 - \beta (1 - q) \Lambda(\mu/N)]^2} \quad (4.100)$$

e finalmente

$$r = \int_0^{\infty} dt \Lambda^2(t) \frac{q}{[1 - \beta (1 - q) \Lambda(t)]^2} \quad (4.101)$$

B - Redes neurais com baixo nível de atividade (76).

Este modelo leva em consideração o fato biológico de que a taxa média de disparos dos neurônios é baixa, bem menor que 50% de uma rede do tipo Hopfield. Durante a maior parte do tempo o neurônio encontra-se inativo.

Dessa forma a rede executa suas tarefas a baixos níveis de atividade e $\langle S_i \rangle$ é muito mais próximo de -1 que zero.

A maneira de tratar matematicamente esse fato é supor que as interações entre os neurônios é expressa por:

$$J_{ij} = \frac{1}{N} \sum_{\mu=1}^P (\xi_i^\mu - a)(\xi_j^\mu - a) \quad (4.102)$$

onde o parâmetro a é o valor médio dos ξ 's e $-1 \leq a \leq 1$.

Considerando que cada ξ_i^μ pode ser escolhido independentemente com probabilidade:

$$P(\xi_i^\mu) = \frac{1}{2}(1+a)\delta(\xi_i^\mu - 1) + \frac{1}{2}(1-a)\delta(\xi_i^\mu + 1) \quad (4.103)$$

definindo a sobreposição média por

$$m^\mu = \frac{1}{N} \sum_{j=1}^N (\xi_j^\mu - a) \langle S_j \rangle \quad (4.104)$$

e o campo médio por

$$\langle h_i \rangle = \langle \sum_{\mu=1}^s (\xi_i^\mu - a) m^\mu \rangle \quad (4.105)$$

Da mesma forma que no item anterior

$$\langle (\delta h_i)^2 \rangle = \langle \left(\sum_{\mu>s}^P (\xi_i^\mu - a) m^\mu \right)^2 \rangle \quad (4.106)$$

Tomando $\alpha r = \langle (\delta h_i)^2 \rangle$ encontramos para os parâmetros de ordem:

$$m^\nu = \langle (\xi^\nu - a) \tanh \beta [\sqrt{\alpha r z} + \sum_{\nu} (m^\mu + h^\nu)(\xi^\mu - a)] \rangle \quad (4.107)$$

$$q = \langle\langle \operatorname{tgh}^2 \beta [\sqrt{\alpha r z} + \sum_{\nu} (m^{\mu} + h^{\nu})(\xi^{\mu} - a)] \rangle\rangle \quad (4.108)$$

O parâmetro r pode ser obtido de:

$$m^{\mu} = \frac{1}{N} \sum_{i=1}^N (\xi_i^{\mu} - a) [\operatorname{tgh}(\beta \sum_{\nu \neq \mu} h_i^{\nu}) + \beta (\xi_i^{\mu} - a) m^{\mu} (1 - \operatorname{tgh}^2(\beta \sum_{\nu \neq \mu} h_i^{\nu}))] \quad (4.109)$$

ou

$$m^{\mu} [1 - \frac{\beta}{N} \sum_{i=1}^N (\xi_i^{\mu} - a) [1 - \operatorname{tgh}^2(\beta \sum_{\nu \neq \mu} h_i^{\nu})]] = \frac{1}{N} \sum_{i=1}^N (\xi_i^{\mu} - a) \operatorname{tgh}(\beta \sum_{\nu \neq \mu} h_i^{\nu}) \quad (4.110)$$

e finalmente,

$$r = \frac{q(1-a^2)}{[1-\beta(1-a^2)(1-q)]^2} \quad (4.111)$$

CAPÍTULO 5.

MODELO PROPOSTO

O objetivo de propormos o modelo descrito a seguir, é unicamente, para aplicação do método dos cumulantes e obtenção de seus parâmetros de ordem. Um estudo mais aprofundado do modelo poderá ser feito posteriormente.

5.1 Cálculo dos parâmetros de ordem do modelo

Neste modelo unidimensional consideraremos uma rede na qual a interação entre os neurônios é dada por:

$$J_{ij} = \frac{1}{N} \sum_{\mu=1}^p \xi_i^{\mu} \xi_j^{\mu} \cos k(x_i - x_j) \quad (5.1)$$

e trataremos apenas o caso em que a capacidade de armazenamento é nula, isto é: o número de padrões armazenados permanece finito quando tomamos o limite termodinâmico.

Um modelo para vidros de spins com interação igual a descrita acima foi abordado nas referências (77-79)

O ponto principal do modelo é levarmos em conta a variação espacial do acoplamento. Evidentemente seria desejável que a interação decrescesse a medida que a distância aumentasse, de forma que um neurônio sofresse pouca ou nenhuma influência de outro que se encontrasse distante, mas tal tratamento envolve grandes complicações matemáticas quando o modelo é não-separável.

A solução, também neste caso, é mais simples quando comparado à obtida pela Mecânica Estatística (APÊNDICE D).

Observamos do modelo que quando $k=0$ recuperamos o modelo de Hopfield.

O campo num dado neurônio i é :

$$h_i = \sum_{j=1}^N J_{ij} \langle S_j \rangle \quad (5.2)$$

ou ainda:

$$h_i = \frac{1}{N} \sum_{j=1}^N \sum_{\mu=1}^p \xi_i^\mu \xi_j^\mu \langle S_j \rangle [\cos kx_i \cos kx_j + \text{sen } kx_i \text{ sen } kx_j] \quad (5.3)$$

Vamos agora assumir uma distribuição uniforme (59) para as variáveis ξ 's. Então, assim o valor médio de $f(x)$ reduz-se a:

$$\langle f(x) \rangle = \int_{-\infty}^{\infty} f(x) P(x) dx = \text{Constante} \cdot f(x) \quad (5.4)$$

Assim usando a equação

$$\langle S_i \rangle = \text{tgh}(\beta h_i) \quad (5.5)$$

e definindo as sobreposições médias por:

$$m^{(1\mu)} = \frac{1}{N} \sum_j \xi_j^\mu \langle S_j \rangle \cos kx_j \quad (5.6)$$

$$m^{(2\mu)} = \frac{1}{N} \sum_j \xi_j^\mu \langle S_j \rangle \text{sen } kx_j \quad (5.7)$$

obtemos que:

$$\langle S_i \rangle = \text{tgh } \beta \left[\sum_\mu \xi_i^\mu (m^{(1\mu)} \cos kx_i + m^{(2\mu)} \text{sen } kx_i) \right] \quad (5.8)$$

que ao ser substituída nas equações das sobreposições acima conduzem a:

$$m^{(1\mu)} = \frac{1}{N} \sum_i \xi_i^\mu \cos kx_i \text{tgh } \beta \left[\sum_\mu \xi_i^\mu (m^{(1\mu)} \cos kx_i + m^{(2\mu)} \text{sen } kx_i) \right] \quad (5.9)$$

$$m^{(2\mu)} = \frac{1}{N} \sum_i \xi_i^\mu \text{sen } kx_i \text{tgh } \beta \left[\sum_\mu \xi_i^\mu (m^{(1\mu)} \cos kx_i + m^{(2\mu)} \text{sen } kx_i) \right] \quad (5.10)$$

Podemos calcular a partir das equações acima, da equação (5.8) e da sobreposição média entre um padrão ensinado e um estado qualquer da rede, isto é, de

$$m^\mu = \frac{1}{N} \sum_j \xi_j^\mu \langle S_j \rangle \quad (5.11)$$

que:

$$m^\mu = \frac{1}{N} \sum_i \xi_i^\mu \text{tgh } \beta \left[\sum_\mu \xi_i^\mu (m^{(1\mu)} \cos kx_i + m^{(2\mu)} \text{sen } kx_i) \right] \quad (5.12)$$

A equação acima descreve a recuperação dos padrões ensinados a rede, isto é, quando $m^\mu \neq 0$ podemos recuperar total ou parcialmente um padrão ensinado.

5.2 Resultados do modelo

No limite termodinâmico, podemos substituir as somatórias por integrais obtendo assim:

$$\frac{1}{N} \sum_{i=1}^N f(k x_i) \rightarrow \frac{1}{2L} \int_{-L}^L f(k x) dx \quad (5.13)$$

Fazendo a transformação de variável

$$x' = \frac{x}{L} \quad (5.14)$$

teremos:

$$\frac{1}{2L} \int_{-L}^L f(k x) dx = \frac{1}{2} \int_{-1}^1 f(k L x') dx' \quad (5.15)$$

Definindo agora $k' = kL$ e considerando que foi ensinado à rede um único padrão, obtemos:

$$\begin{bmatrix} m^1(t+1) \\ m^2(t+1) \end{bmatrix} = \begin{bmatrix} \frac{1}{2} \int_{-1}^1 \cos(kx) \tanh[\beta (m^1(t) \cos(kx) + m^2(t) \sin(kx))] dx \\ \frac{1}{2} \int_{-1}^1 \sin(kx) \tanh[\beta (m^1(t) \cos(kx) + m^2(t) \sin(kx))] dx \end{bmatrix} \quad (5.16)$$

Após os valores dos parâmetros m^1 e m^2 convergirem, calculamos então o parâmetro m .

Os resultados observados são que:

Quando $T > 1$ todos os parâmetros se anulam quaisquer que sejam os valores de k' . (Fase paramagnética)

Para $k=0$, m assume seu valor máximo, aproximando-se do valor 1 a medida que a temperatura decresce pois quanto menor o nível de ruído da rede, maior sua capacidade em recuperar os padrões ensinados.

Os gráficos abaixo, tomando as temperaturas $T=0.1$, $T=0.5$, e $T=0.55$ ilustram os resultados descritos acima.

Observação:

As valores das abscissas de todos os gráficos deste capítulo são múltiplos de π

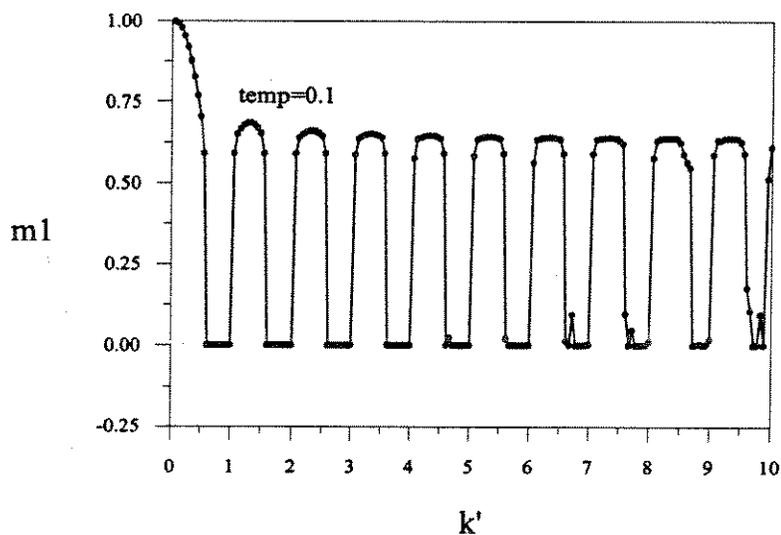


Figura 13 - Gráfico de m^1 versus k' .

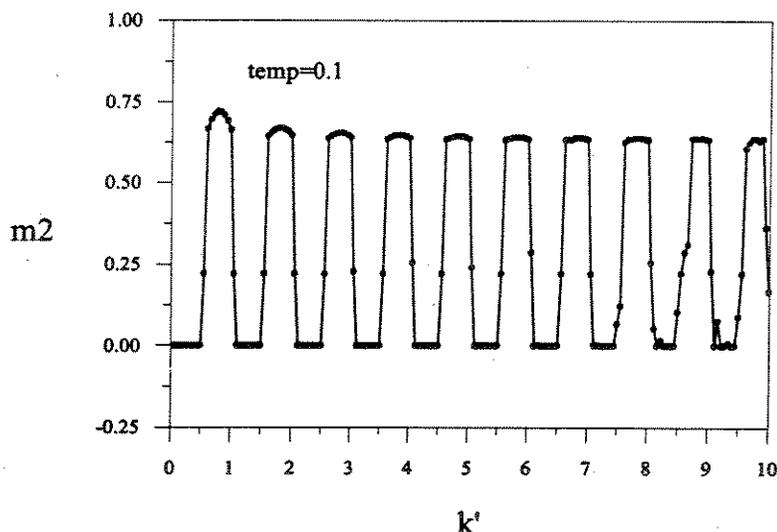


Figura 14 - Gráfico de m^2 versus k' .

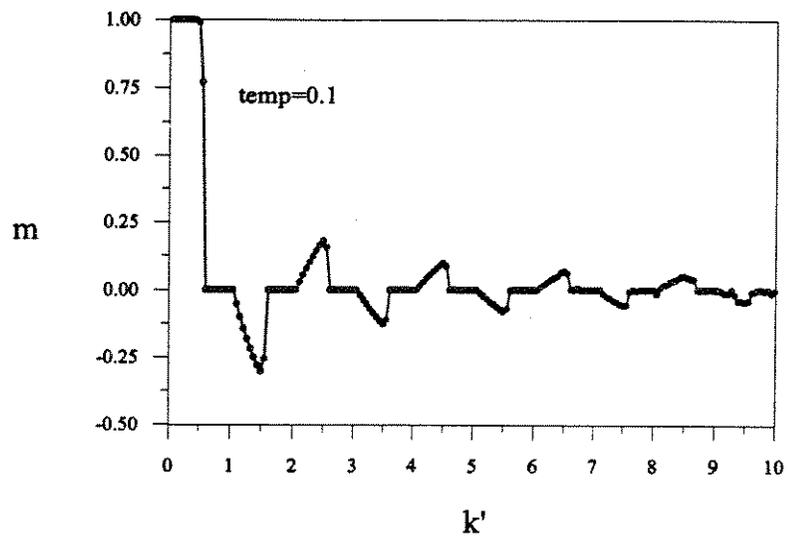


Figura 15 - Gráfico de m versus k' .

A FIGURA 15, mostra-nos que a baixas temperaturas, a rede consegue recuperação total dos padrões para valores de k' menores que $\pi/2$. Acima deste valor o parâmetro m torna-se muito pequeno, as vezes assumindo valores negativos (ocorre a recuperação do inverso do padrão ensinado).

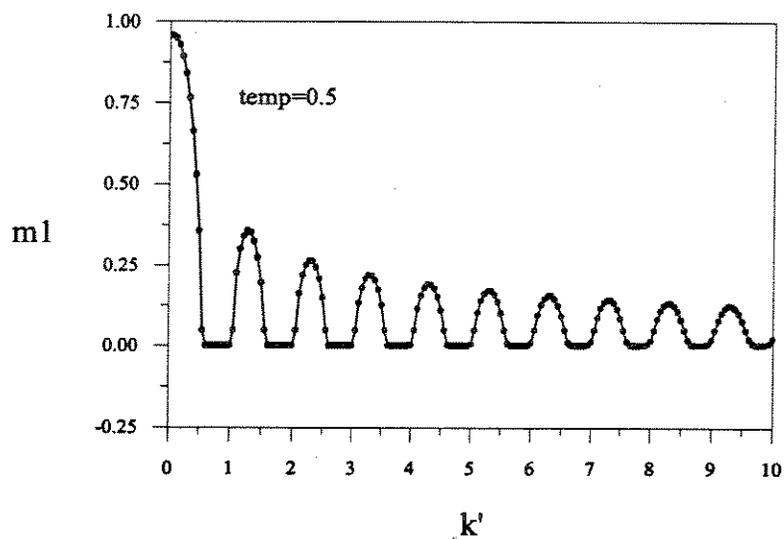


Figura 16 - Gráfico de m^1 versus k' .

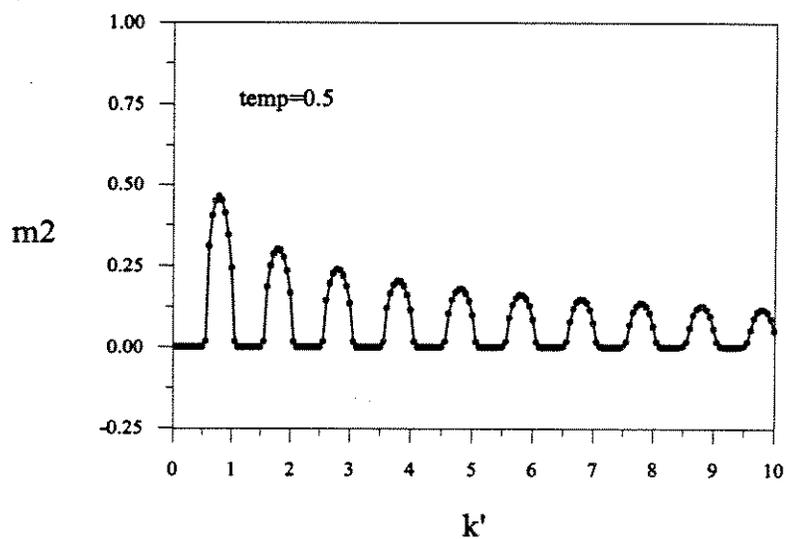


Figura 17 - Gráfico de m^2 versus k' .

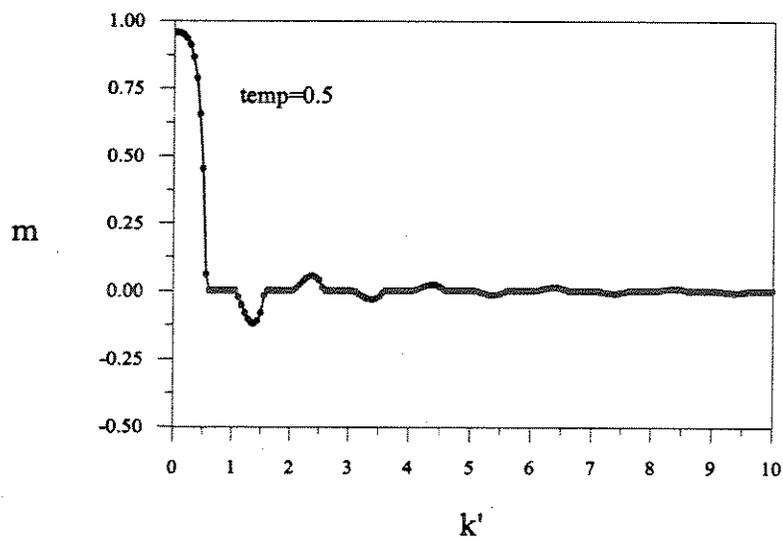


Figura 18 - Gráfico de m versus k' .

Da FIGURA 18 observamos que ao aumentarmos o nível de ruído, houve uma diminuição do valor de m . Observamos ainda que para valores de k' menores que $\pi/2$ ocorre a recuperação do padrão ensinado

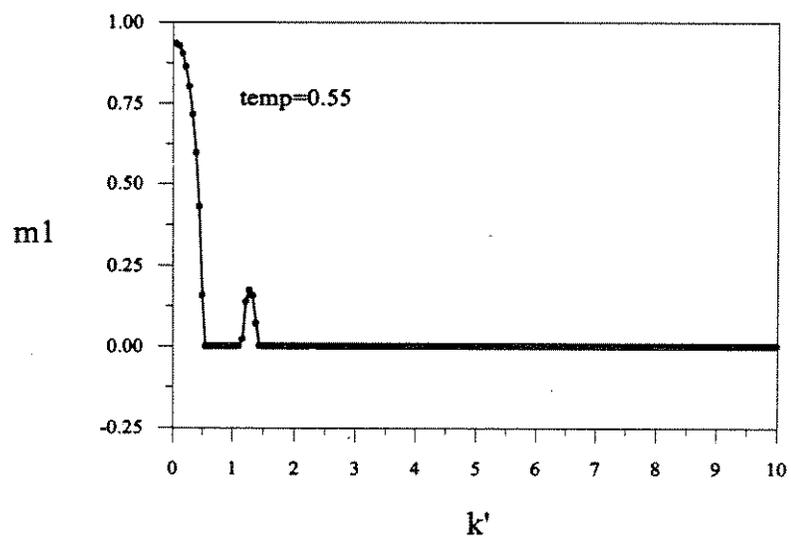


Figura 19 - Gráfico de m^1 versus k' .

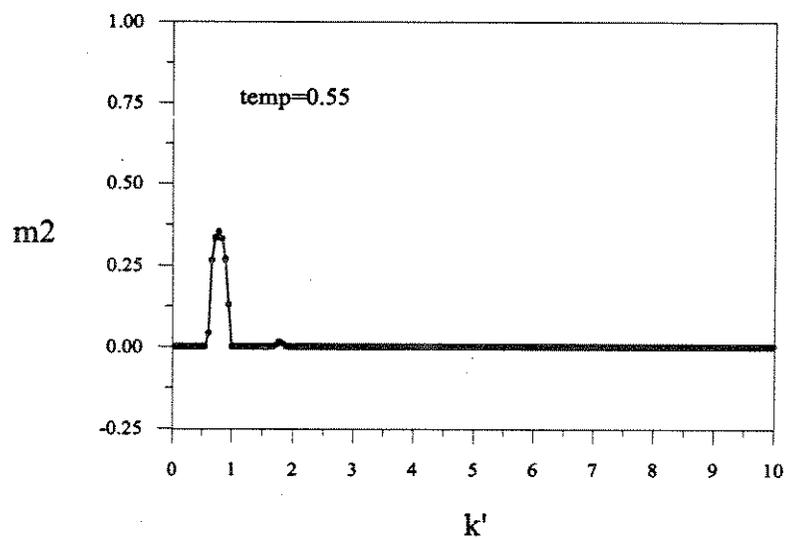


Figura 20 - Gráfico de m^2 versus k' .

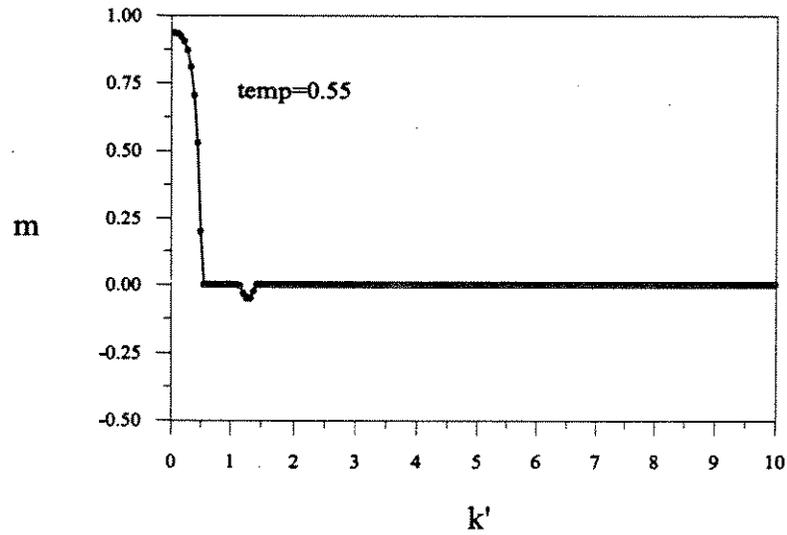


Figura 21 - Gráfico de m versus k'.

E finalmente da FIGURA 21, observamos a diminuição do valor de m em $k'=0$ e que a recuperação do padrão ensinado à rede ocorre apenas para valores de k' menores que $\pi/2$.

As simulações efetuadas mostram que há recuperação apenas para valores de k tais que:

$$k \leq \frac{\pi}{2L} \quad (5.17)$$

O resultado acima pode também ser obtido através de uma análise mais profunda sobre as equações do modelo.

CAPÍTULO 6

CONCLUSÕES

Mostrou-se neste trabalho que podemos encontrar os parâmetros de ordem de redes neurais, usando o método dos cumulantes para vários tipos de redes, de forma mais simples que os métodos da Mecânica Estatística. Primeiramente, tratamos o modelo de Hopfield, a seguir, um modelo de rede neural com esquecimento e também um modelo de rede com baixo nível de atividade neural.

Em seguida, aplicamos o método a um modelo no qual, a interação espacial entre os neurônios da rede varia de forma cossenoidal. Verificamos que, em tal tipo de rede, o aumento da frequência de oscilação, acima de $\frac{\pi}{2L}$, faz com que a rede não possa mais recuperar os padrões ensinados.

Trataremos nessas considerações finais, de outros modelos que, devido a sua semelhança aos modelos já abordados, poderão ser estudados através do método dos cumulantes.

Um destes problemas é o da bipartição de grafos. Uma aplicação consiste no projeto de circuitos integrados, no qual, N elementos devem ser colocados numa única pastilha. Se isto não for possível, gostaríamos que metade de tais elementos ficasse numa pastilha, e metade em outra de forma que o número de conexões (fios) entre as pastilhas fosse o menor possível. Escolher quais elementos colocar numa pastilha é um problema de otimização, com a função custo sendo o número de fios conectando os circuitos. Este problema é abordado nas referências (125-130). Numa

análise preliminar, obteve-se através do método dos cumulantes, os mesmos resultados da referência (125), que faz uso do método das réplicas.

O método também poderá ser aplicado a vários outros modelos de redes neurais como por exemplo aos modelos propostos por Personnaz e outros (131), Fontanari (132), Kanter e outros (133), Gutfreund (134) e Kühn (135) que utilizam-se do método das réplicas para a obtenção dos parâmetros de ordem.

Evidentemente o método das réplicas é muito mais rico conceitualmente e mais poderoso em seus resultados, mas sofre ainda de alguns problemas não totalmente resolvidos, principalmente ligados ao fato do limite do número de réplicas do sistema $n \rightarrow 0$, entropia negativa, ou ainda matrizes de dimensão nula (136-138).

Com o método dos cumulantes obtêm-se os resultados de forma mais rápida quando considerarmos simetria das réplicas e quando a distribuição das variáveis envolvidas for gaussiana (que felizmente é a grande maioria dos casos).

No caso de outras distribuições, como por exemplo, o problema do caixeiro viajante (139-143) o método dos cumulantes deve ser de difícil aplicação, assim como o método das réplicas.

Observa-se, que o método dos cumulantes, tem uma grande aplicação em áreas da Física que tratam de sistemas de partículas interagindo, quando as propriedades a serem calculadas, são expressas por funções exponenciais (144-146). Para um conjunto de partículas interagindo, os cumulantes podem ser usados a fim de obtermos uma expressão aproximada da energia livre, conduzindo a cálculos mais simples e diretos que os tradicionais.

Pretendemos, como desenvolvimento futuro, aplicar o método a um modelo em que a interação entre os neurônios decresça espacialmente, numa tentativa de simular o modelo de Kohonen (147) ("chapéu mexicano") no qual um grupo de neurônios ao disparar seu potencial de ação, inibe seus vizinhos. Tal tipo de interação poderia ser descrito por :

$$J_{ij} = \frac{1}{N} \sum_{\mu=1}^p \xi_i^{\mu} \xi_j^{\mu} e^{-\alpha(x_i - x_j)} \cos k(x_i - x_j) \quad (6.1)$$

Um modelo com interação exponencial foi estudado para o caso de vidros de spins por Almeida (148).

Espera-se então, que este método sirva de uma ferramenta prática a outros profissionais não treinados em Mecânica Estatística de sistemas desordenados, que desejam um conhecimento mais profundo sobre redes neurais.

APÊNDICE A.

VIDROS DE SPINS

Existe uma grande analogia entre uma rede neural do tipo Hopfield e os vidros de spins. Para um estudo mais completo, sugerimos as referências (80-90).

Durante as décadas de 50 e 60, as ligas magnéticas onde átomos com momentos magnéticos (Fe, Mn, Cr) estão combinados aleatoriamente com metais não-magnéticos (Au, Ag, Cu) foram bastante estudadas objetivando entender a natureza da estrutura amorfa e o magnetismo resultante, bem como investigar fenômenos críticos e percolação.

Destes estudos surgiu o termo "vidro de spin" (spin glass) creditado por alguns autores a Brian R. Coles (91,92), embora haja controvérsias sobre a origem do termo (93). Ele usou o termo para descrever a classe de ligas magnéticas moderadamente diluídas onde, da mesma maneira que os átomos em um vidro ordinário estão distribuídos aleatoriamente sem nenhuma estrutura regular, também os spins dos átomos magnéticos nestas ligas estão diluídos aleatoriamente em uma matriz de átomos não magnéticos.

O comportamento destas ligas tem mostrado que os átomos magnéticos estão muito afastados para possuírem a estrutura de um metal e suficientemente próximos para que as interações de troca dominem sobre outros tipos de interações como o efeito Kondo e aquelas devidas aos elétrons livres.

Em 1971, Cannella, V. Mydosh, J.A. e Budnick, J.I. (94) medindo a susceptibilidade magnética das ligas AuFe (1% a 8% em Fe) observaram uma mudança brusca na susceptibilidade a uma temperatura T_c . Foi somente após a observação deste pico na susceptibilidade magnética numa temperatura bem definida que conjecturou-se sobre a possibilidade da ocorrência de uma transição de fase termodinâmica, assunto ainda controverso uma vez que, deveria ser verificada uma mudança abrupta numa propriedade termodinâmica da liga, tal como no calor específico. Entretanto, o calor específico de tipos diferentes de ligas que apresentam a fase vidro de spin, mostram uma curva contínua, com um máximo ocorrendo a uma temperatura aproximadamente 20% maior que a temperatura de transição T_c .

No interior de um vidro de spin, os átomos formam uma rede não-magnética na qual estão diluídos os átomos do material que possui momento magnético. Nesta faixa de diluição, e abaixo de T_c a interação predominante é a interação de troca e a constante de acoplamento entre os spins dos átomos de ferro, J_{ij} , é do tipo RKKY (95-97) na qual J_{ij} , é função oscilatória da distância r entre os átomos e é expressa por:

$$J(r_{ij}) = A \frac{\cos(kr_{ij})}{(kr_{ij})^3} \quad (\text{A.1})$$

Em tal tipo de interação, observamos que dependendo da distância entre os átomos de ferro teremos $J_{ij} > 0$ (interações ferromagnéticas) e $J_{ij} < 0$ (interações antiferromagnéticas). Desta forma num vidro de spin teremos aproximadamente, metade dos átomos interagindo ferromagneticamente e a outra metade antiferromagneticamente. Em tal situação, uma grande parte dos átomos do material que possuem momento magnético não consegue satisfazer suas interações com o restante dos átomos de ferro da rede ocorrendo o que chamamos de frustração (98). A frustração entre as interações, causa o aparecimento de um grande número de estados, todos com a mesma energia, que é uma característica dos vidros de spin onde surge um novo tipo de ordem magnética.

Um dos resultados importantes dos vidros de spins foi a introdução por Edwards e Anderson (71) de um parâmetro de ordem temporal definido por:

$$q = \lim_{t \rightarrow \infty} \langle S_i(0)S_i(t) \rangle \quad (\text{A.2})$$

cujo conjunto de chaves interno representa a média térmica e o conjunto externo representa a média sobre todos os spins.

A definição de q advém do fato de num vidro de spin estarem ausentes os parâmetros convencionais os quais representam correlações espaciais de longo alcance.

Para interpretarmos fisicamente o parâmetro de ordem q é importante observarmos a diferença entre as fases paramagnética e de vidro de spin.

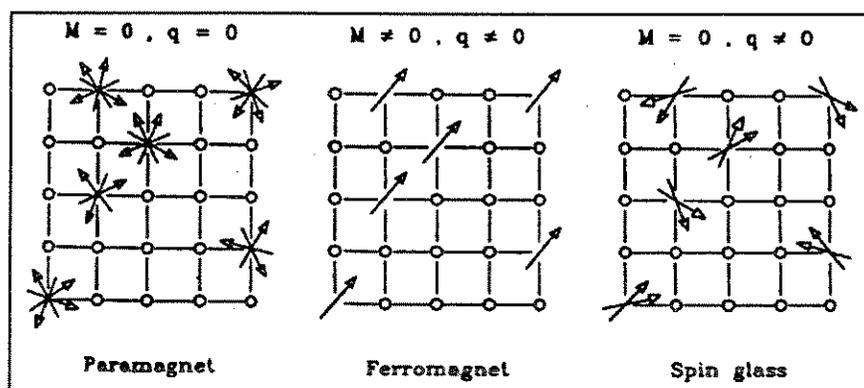


Figura A.1 - Fases Magnéticas

Na fase paramagnética, quando $T > T_c$, os spins dos átomos estão totalmente desordenados, isso implica que a magnetização é nula, $m = 0$ e $q = 0$. Abaixo da temperatura crítica, na fase de vidro de spin os mesmos foram "congelados" (em média) em direções fixas, resultando $q \neq 0$ e $m = 0$. Isto é devido ao fato de que se fosse possível obtermos uma fotografia da fase paramagnética e também da de vidro de spin num instante $t = 0$ nós não conseguiríamos distinguir uma fase da outra, pois em ambas, os spins encontram-se desordenados apontando em todas as direções. Mas se obtivéssemos uma segunda fotografia num tempo $t > 0$, observaríamos que na fase de vidro de spin, os mesmos conservariam em média as mesmas orientações enquanto na fase paramagnética eles teriam mudado de direção. Assim podemos afirmar que na fase de vidro de spin existe uma espécie de "memória".

O principal modelo dos vidros de spins é o modelo de Sherrington-Kirkpatrick (modelo SK) (99)

APÊNDICE B.

O MÉTODO DAS RÉPLICAS

Várias das propriedades termodinâmicas de um sistema são obtidas através da sua energia livre F , (100-102), que pode ser calculada a partir da função de partição Z . A fim de calcularmos a energia livre devemos calcular a média do $\langle \ln Z \rangle$ sobre a distribuição $P(J_{ij})$, ou seja:

$$f = \langle f[J] \rangle = -\frac{kT}{N} \langle \ln Z[J] \rangle \quad (\text{B.1})$$

Obviamente não é tarefa fácil, (82) pois necessitamos calcular a energia livre, a qual depende de um número muito grande das variáveis (J) que descrevem a desordem.

A média da equação (B.1) não pode ser feita diretamente devido ao fato de que as variáveis aleatórias (J) encontrarem-se como argumento da função logarítmica.

Utilizamos desta forma, um artifício matemático conhecido como teoria das réplicas empregado pela primeira vez na mecânica estatística por Lin e Kac (103) e redescoberto independentemente por Edwards (71), Grinstein e Luther (104) e Emery (105)

O procedimento consiste em usar a seguinte identidade matemática:

$$\ln x = \lim_{n \rightarrow 0} \frac{(x^n - 1)}{n} \quad (\text{B.2})$$

Para n inteiro e positivo podemos expressar $Z^n[J]$ em termos de n réplicas idênticas ao sistema por:

$$Z^n[J] = \prod_{\alpha=1}^n Z_{\alpha}[J] = \prod_{\alpha=1}^n \exp[-\beta H(J, S_i^{\alpha})] \quad (\text{B.3})$$

ou ainda

$$Z^n[J] = \exp[-\beta \sum_{\alpha=1}^n H(J, S_i^{\alpha})] \quad (\text{B.4})$$

onde Z_{α} é a função partição da α -ésima réplica. O índice α varia de $1, \dots, n$ e onde cada valor pode ser interpretado como uma réplica do sistema original. Para calcularmos a energia livre do sistema original utilizamos então

$$f = -\frac{1}{\beta} \lim_{N \rightarrow \infty} \lim_{n \rightarrow 0} N^{-1} (\langle Z^n \rangle - 1) \quad (\text{B.5})$$

A interação entre duas réplicas α e β pode ser interpretada como representando o sistema em dois instantes distintos de tempo. Assim, a correlação entre duas réplicas do sistema comporta-se de forma idêntica à correlação que define o parâmetro de ordem de Edwards-Anderson. Então podemos calcular q por:

$$q_{\alpha\beta} = \lim_{t \rightarrow \infty} \langle\langle S_i^{\alpha}(0) S_i^{\beta}(t) \rangle\rangle \quad (\text{B.6})$$

APÊNDICE C

APLICAÇÃO DOS MÉTODOS DA MECÂNICA ESTATÍSTICA AO MODELO DE HOPFIELD

C.1 Caso $\alpha=0$ ($\alpha=p/N$)

Vamos tratar o modelo de Hopfield com neurônios estocásticos com a regra de aprendizado generalizada de Hebb, primeiramente no caso em que o número de padrões armazenados $p \ll N$. O trabalho foi desenvolvido por Amit, Gutfreund e Sompolinski (AGS-1) em 1985 (106).

Seja

$$H = -\frac{1}{2} \sum_{i \neq j}^N J_{ij} S_i S_j \quad (\text{C.1})$$

$$J_{ij} = \frac{1}{N} \sum_{\mu=1}^p \xi_i^{\mu} \xi_j^{\mu} \quad (\text{C.2})$$

Substituindo J_{ij} no Hamiltoniano obtemos:

$$H_0 = -\frac{1}{2N} \sum_{\mu=1}^p \left(\sum_{i=1}^N \xi_i^\mu S_i \right)^2 + \frac{p}{2} \quad (\text{C.3})$$

Seja

$$\alpha = \frac{p}{N} \quad (\text{C.4})$$

a razão do número de memórias armazenadas pelo número de neurônios da rede. Consideremos o sistema no limite termodinâmico ($N \rightarrow \infty$).

Vamos adicionar a H_0 um conjunto de campos externos $h^\mu \xi_i^\mu$, um para cada padrão ξ_i^μ . (À frente faremos $h^\mu = 0$ após utilizarmos estes campos). Então:

$$H = H_0 - \sum_{\mu=1}^p h^\mu \sum_{i=1}^N \xi_i^\mu S_i \quad (\text{C.5})$$

e a função partição fica:

$$Z = e^{-\frac{\beta p}{2}} \text{Tr}_S \exp \left[\frac{\beta}{2N} \sum_{\mu=1}^p \left(\sum_{i=1}^N \xi_i^\mu S_i \right)^2 + \beta \sum_{\mu=1}^p h^\mu \sum_{i=1}^N \xi_i^\mu S_i \right] \quad (\text{C.6})$$

Utilizando-se da propriedade das integrais gaussianas (107,108), transformamos a parte quadrática da função partição em linear com o custo de introduzirmos a variável auxiliar x e sua integral. Como existem p termos quadráticos na função partição introduzimos p variáveis auxiliares m^μ e fazendo

$$a = \frac{\beta N}{2} \quad \text{e} \quad b^\mu = \beta \sum_{i=1}^N \xi_i^\mu S_i \quad (\text{C.7})$$

obtemos:

$$Z = e^{-\frac{\beta p}{2} \left(\frac{\beta N}{2\pi} \right)^{\frac{p}{2}}} \text{Tr}_S \prod_{\mu=1}^p \int dm^\mu \exp \left(-\frac{1}{2} \beta N (m^\mu)^2 + \beta (m^\mu + h^\mu) \sum_{i=1}^N \xi_i^\mu S_i \right) \quad (\text{C.8})$$

Vamos adotar a notação vetorial de $\vec{m}, \vec{h}, \vec{\xi}_i$ como vetores de p componentes m^μ, h^μ, ξ_i^μ respectivamente. Então:

$$Z = e^{-\frac{\beta p}{2} \left(\frac{\beta N}{2\pi} \right)^{\frac{p}{2}}} \int d\vec{m} e^{-\frac{\beta N \vec{m}^2}{2}} \prod_{i=1}^N \text{Tr}_{S_i} e^{\beta (\vec{m} + \vec{h}) \cdot \vec{\xi}_i} \quad (\text{C.9})$$

O traço pode agora ser facilmente calculado pois a exponencial é linear em S_i , assim usando

$$e^x + e^{-x} = 2 \cosh(x) \quad (\text{C.10})$$

obtemos após reorganizarmos a função partição:

$$Z = \left(\frac{\beta N}{2\pi} \right)^{\frac{p}{2}} \int d\vec{m} e^{-\beta N f(\beta, \vec{m})} \quad (\text{C.11})$$

com

$$f(\beta, \vec{m}) = \frac{1}{2} \alpha + \frac{1}{2} \vec{m}^2 - \left\langle \frac{1}{\beta N} \sum_{i=1}^N \ln [2 \cosh \{ \beta (\vec{m} + \vec{h}) \cdot \vec{\xi}_i \}] \right\rangle \quad (\text{C.12})$$

O fato de que na integral acima, o expoente é proporcional a N permite avaliar esta integral quando $N \rightarrow \infty$. Quanto maior N , mais a integral é dominada por contribuições advindas de regiões onde $f(\beta, \vec{m})$ é menor. Assim, podemos aproximá-la por encontrar o valor de \vec{m} que minimiza a função e expandir o integrando em torno deste ponto. Esse método é conhecido como método do ponto de sela (109,110) (saddle point method).

Seja por exemplo a integral:

$$I = \sqrt{N} \int dx e^{-Ng(x)} \quad (\text{C.13})$$

Expandindo o expoente em torno de x_0 que supomos ser o mínimo de $g(x)$ obtemos:

$$I = \sqrt{N} \int dx \exp \left[-N \left\{ g(x_0) + \frac{1}{2} g''(x_0)(x-x_0)^2 + \dots \right\} \right] \quad [g'(x_0) = 0] \quad (C.14)$$

$$I = \sqrt{N} e^{-Ng(x_0)} \int dx \exp \left[-\frac{N}{2} g''(x_0)(x-x_0)^2 \right] \quad (C.15)$$

Observamos que a integral acima é gaussiana e calculando-a obtemos:

$$I = e^{-Ng(x_0)} \sqrt{\frac{2\pi}{g''(x_0)}} \quad (C.16)$$

Para $N \rightarrow \infty$ o resultado é dominado pelo fator exponencial como podemos observar de:

$$-\frac{1}{N} \ln I = g(x_0) + \frac{1}{2N} (\ln g''(x_0) - \ln 2\pi) \quad (C.17)$$

Então, para $N \rightarrow \infty$

$$-\frac{1}{N} \ln I \rightarrow g(x_0) \quad (C.18)$$

Assim tudo o que necessitamos é encontrar x_0 , denominado ponto de sela. Então obtemos para a integral em (C.13)

$$-\frac{1}{N} \ln Z = \beta \min_{\bar{m}} f(\beta, \bar{m}) \quad N \rightarrow \infty \quad (C.19)$$

Como

$$-\frac{1}{N} \ln Z = \frac{F}{N} \quad (\text{C.20})$$

temos que f é a energia livre por neurônio.

Para minimizarmos $f(\beta, \bar{m})$ efetuamos:

$$\frac{\partial f}{\partial m^\mu} = 0 = m^\mu - \frac{1}{N} \sum_{i=1}^N \xi_i^\mu \operatorname{tgh}[\beta (\bar{m} + \bar{h}) \cdot \bar{\xi}_i] \quad (\text{C.21})$$

Obtemos então um conjunto de p equações simultâneas não lineares para os m^μ expressas por:

$$m^\mu = \frac{1}{N} \sum_{i=1}^N \xi_i^\mu \operatorname{tgh}[\beta (\bar{m} + \bar{h}) \cdot \bar{\xi}_i] \quad \mu = 1, \dots, p \quad (\text{C.22})$$

Na equação acima e na da energia livre (C.12), podemos substituir a média $\frac{1}{N} \sum_{i=1}^N$ sobre os componentes dos padrões, por uma média sobre os padrões. Um sistema com esta propriedade é denominado de auto-mediado (self-averaging). Essa propriedade advém do fato de que quando fazemos $\frac{1}{N} \sum_{i=1}^N$ estamos escolhendo N ξ_i 's independentes de uma distribuição $P(\bar{\xi})$ a qual supomos ser uniforme sobre as 2^p possibilidades. Assim se N é muito grande comparado a 2^p , então a média sobre os componentes de ξ_i^μ 's é equivalente a uma média sobre a distribuição dos ξ_i 's.

Então as equações (C.12) e (C.22) com $\alpha \rightarrow 0$ resultam:

$$m^\mu = \langle\langle \xi_i^\mu \operatorname{tgh}[\beta (\bar{m} + \bar{h}) \cdot \bar{\xi}_i] \rangle\rangle \quad (\text{C.23})$$

$$f(\beta, \bar{m}) = \frac{1}{2} \bar{m}^2 - \beta^{-1} \langle\langle \ln \{2 \cosh[\beta (\bar{m} + \bar{h}) \cdot \bar{\xi}_i]\} \rangle\rangle \quad (\text{C.24})$$

onde $\langle\langle \dots \rangle\rangle$ indica uma média sobre a distribuição dos padrões.

Os valores de m^μ admitem uma interpretação física. Se diferenciarmos a energia livre ($F = -\beta^{-1} \ln Z$) em relação a h^μ obtemos:

$$\frac{\partial F}{\partial h^\mu} = -\beta^{-1} \frac{\partial \ln Z}{\partial h^\mu} = -\sum_{i=1}^N \xi_i^\mu \langle S_i \rangle \quad (C.25)$$

Por outro lado das equações (C.21) e (C.25) obtemos:

$$\frac{\partial F}{\partial h^\mu} = N \frac{\partial f}{\partial h^\mu} = -N \langle \xi^\mu \operatorname{tgh}[\beta (\bar{m} + \bar{h}) \cdot \bar{\xi}_i] \rangle = -Nm^\mu \quad (C.26)$$

e podemos então identificar

$$m^\mu = \frac{1}{N} \sum_{i=1}^N \xi_i^\mu \langle S_i \rangle \quad (C.27)$$

assim, o valor de m^μ , que é o ponto de sela, é justamente o produto escalar entre o padrão ξ^μ e a configuração média da rede.

A fim de derivarmos a equação anterior, fizemos uso dos campos externos h^μ . Supondo agora que estes campos externos sejam nulos resultará:

$$m^\mu = \langle \xi^\mu \operatorname{tgh}(\beta \bar{m} \cdot \bar{\xi}_i) \rangle \quad (C.28)$$

C.2 Caso $\alpha \neq 0$ (AGS-2) (74,111-124).

A quantidade básica a ser calculada é novamente $\ln Z$. Mas Z depende agora de um conjunto particular de padrões usado para calcular os J_{ij} , através da regra de Hebb. Devemos então, calcular a média $\langle \ln Z \rangle$ sobre a distribuição de todos os padrões possíveis. Assim, obteremos a energia livre média, cujas derivadas serão as quantidades que desejamos, tais como os m^μ . Como já vimos (vidros de spins), esta média, que é diferente de $\ln \langle Z \rangle$ é extremamente difícil (se não impossível) de calcular.

Utilizamos do método das réplicas que consiste em utilizarmos:

$$\langle\langle \ln Z \rangle\rangle = \lim_{n \rightarrow 0} \frac{\langle\langle Z^n \rangle\rangle - 1}{n} \quad (C.29)$$

que permitirá calcular $\ln \langle\langle Z \rangle\rangle$ a partir de $\langle\langle Z^n \rangle\rangle$. Para o cálculo de $\langle\langle Z^n \rangle\rangle$ usamos n réplicas do sistema original, de forma que:

$$Z^n = \text{Tr}_{S^1} \text{Tr}_{S^2} \dots \text{Tr}_{S^n} \exp[-\beta(E\{S_i^1\} + \dots + E\{S_i^n\})] \quad (C.30)$$

Cada cópia do sistema original é denotada por um índice em seus componentes S_i^j 's variando de 1 a n .

Então como no caso $\alpha = 0$

$$\begin{aligned} \langle\langle Z^n \rangle\rangle = e^{-\frac{\beta p n}{2}} \text{Tr}_{S^p} \langle\langle \int \prod_{\mu p} \frac{dm_p^\mu}{\sqrt{2\pi}} \exp[\beta N(-\frac{1}{2} \sum_{\mu p} (m_p^\mu)^2 + \sum_{\mu p} m_p^\mu \frac{1}{N} \sum_i \xi_i^\mu S_i^p)] \\ \int \prod_{\nu p} \frac{dm_p^\nu}{\sqrt{2\pi}} \exp\{\beta N[-\frac{1}{2} \sum_{\nu p} (m_p^\nu)^2 + \sum_{\nu p} (m_p^\nu + h^\nu) \frac{1}{N} \sum_i \xi_i^\nu S_i^p]\} \rangle\rangle \end{aligned} \quad (C.31)$$

onde p indica as n diferentes réplicas.

Quando o número de padrões armazenados tende ao infinito existe a possibilidade de que a sobreposição do vetor de estado da rede com cada padrão tenda a zero. A fim de excluir este tipo de solução sem muito interesse, assumimos que o vetor de estado possui sobreposição com um número finito s dos p padrões armazenados que serão denominados de padrões condensados. Tal fato permitirá que o sistema seja automeediato e os m_p^ν 's assumam valores apreciáveis para os ν padrões variando de 1 a s . Já para os μ padrões variando de $s+1$ a p , assumimos que os $m_p^\mu \ll 1$.

Vamos agora considerar a contribuição dos μ padrões na fórmula acima denominando a exponencial da primeira integral de A . Então:

$$A = \exp[-\frac{1}{2} \sum_{\mu p} (m_p^\mu)^2 + \frac{\beta}{2N} \sum_{\mu p \sigma} m_p^\mu S_i^p m_p^\mu S_i^\sigma] \quad (C.32)$$

onde as variáveis m_p^μ foram rescalonadas para $m_p^\mu = m_p^\mu \sqrt{\beta N}$ e definindo os elementos de uma matriz K por:

$$k_{\rho\sigma} = \delta_{\rho\sigma} - \frac{\beta}{N} \sum_i S_i^\rho S_i^\sigma \quad (\text{C.33})$$

então A pode ser escrito como:

$$A = \exp\left[-\frac{\beta N}{2} \sum_{\mu\rho\sigma} k_{\rho\sigma} m_p^\mu m_\sigma^\mu\right] \quad (\text{C.34})$$

e utilizando as propriedades das integrais gaussianas (59) obtemos:

$$\int \prod_{\mu\rho} \frac{dm_p^\mu}{\sqrt{2\pi}} \exp\left[-\frac{\beta N}{2} \sum_{\mu\rho\sigma} k_{\rho\sigma} m_p^\mu m_\sigma^\mu\right] = [\det K]^{-\frac{P}{2}} = \exp\left[-\frac{P}{2} \text{Tr}(\ln K)\right] \quad (\text{C.35})$$

Por outro lado se definirmos os elementos de uma matriz Q como:

$$q_{\rho\sigma} = (1 - \delta_{\rho\sigma}) \frac{1}{N} \sum_i S_i^\rho S_i^\sigma \quad (\text{C.36})$$

então a matriz K pode ser escrita como:

$$K = (1 - \beta)I - \beta Q \quad (\text{C.37})$$

Se agora utilizarmos que:

$$e^{Ng(x)} = \int dq e^{Ng(q)} \delta(q - x) = \frac{1}{2\pi i} \int_{-i\infty}^{i\infty} dr \int dq e^{Ng(q)} e^{-Nr(q-x)} = \frac{1}{2\pi i} \int_{-i\infty}^{i\infty} dr \int dq e^{N(g(q) - r(q-x))} \quad (\text{C.38})$$

onde utilizamos a função delta

$$\delta(x) = \frac{1}{2\pi i} \int_{-i\infty}^{i\infty} dr e^{-rx} \quad (C.39)$$

então como existem $n(n-1)/2$ integrais para $r_{\rho\sigma}$ e para $q_{\rho\sigma}$ e como $q_{\rho\sigma} = q_{\sigma\rho}$ e $q_{\rho\rho} = 0$ obtemos:

$$e^{-\frac{1}{2}p \text{Tr}_{S^p} \ln K} = \int \prod_{(\rho\sigma)} dr_{\rho\sigma} \int \prod_{(\rho\sigma)} dq_{\rho\sigma} \exp\left[-\frac{1}{2}p \text{Tr}(\ln K)\right] \exp\left[-\frac{1}{2}N\alpha\beta^2 \sum_{\rho\sigma} r_{\rho\sigma} q_{\rho\sigma} + \frac{1}{2}\alpha\beta^2 \sum_{i\rho\sigma} r_{\rho\sigma} S_i^\rho S_i^\sigma\right] \quad (C.40)$$

onde a notação $(\rho\sigma)$ significa todos os pares diferentes, então

$$\begin{aligned} \langle\langle Z^n \rangle\rangle &= e^{-\frac{\beta pn}{2}} \int \prod_{(\rho\sigma)} dr_{\rho\sigma} \int \prod_{(\rho\sigma)} dq_{\rho\sigma} \exp\left\{N\left[-\frac{1}{2}\alpha \text{Tr}(\ln K) - \frac{1}{2}\alpha\beta^2 \sum_{\rho\sigma} r_{\rho\sigma} q_{\rho\sigma}\right]\right\} \\ \text{Tr}_{S^p} \langle\langle \int \prod_{\nu\rho} \frac{dm_\rho^\nu}{\sqrt{2\pi}} \exp\left\{\beta N\left[-\frac{1}{2}\sum_{\nu\rho} (m_\rho^\nu)^2 + \sum_{\nu\rho} (m_\rho^\nu + h^\nu) \frac{1}{N} \sum_i \xi_i^\nu S_i^\rho\right]\right\} \exp\left[\frac{1}{2}\alpha\beta^2 \sum_{i\rho\sigma} r_{\rho\sigma} S_i^\rho S_i^\sigma\right] \rangle\rangle & \end{aligned} \quad (C.41)$$

ou ainda

$$\begin{aligned} \langle\langle Z^n \rangle\rangle &= e^{-\frac{\beta pn}{2}} \int \prod_{\nu\rho} \frac{dm_\rho^\nu}{\sqrt{2\pi}} \int \prod_{(\rho\sigma)} dr_{\rho\sigma} \int \prod_{(\rho\sigma)} dq_{\rho\sigma} \exp\left\{N\left[-\frac{1}{2}\alpha \text{Tr}(\ln K) - \frac{1}{2}\alpha\beta^2 \sum_{\rho\sigma} r_{\rho\sigma} q_{\rho\sigma} - \frac{1}{2}\beta \sum_{\nu\rho} (m_\rho^\nu)^2\right]\right\} \\ \text{Tr}_{S^p} \langle\langle \exp\left\{\beta\left[\frac{1}{2}\alpha\beta^2 \sum_{i\rho\sigma} r_{\rho\sigma} S_i^\rho S_i^\sigma + \sum_{\nu\rho} (m_\rho^\nu + h^\nu) \frac{1}{N} \sum_i \xi_i^\nu S_i^\rho\right]\right\} \rangle\rangle & \end{aligned} \quad (C.42)$$

A expressão compreendida pelo traço sobre as variáveis S_i^ρ pode agora ser calculada como:

$$\langle\langle \exp \sum_i \ln \left\{ \text{Tr}_{S^p} \exp \left\{ \beta \left[\frac{1}{2} \alpha \beta^2 \sum_{\rho\sigma} r_{\rho\sigma} S^\rho S^\sigma + \sum_{\nu\rho} (m_\rho^\nu + h^\nu) \xi_i^\nu S^\rho \right] \right\} \right\} \rangle\rangle \quad (C.43)$$

ou

$$\exp\{N \ll \ln \{ \text{Tr}_{S^p} \exp \beta [\frac{1}{2} \alpha \beta^2 \sum_{\rho \neq \sigma} r_{\rho\sigma} S^\rho S^\sigma + \sum_{\nu p} (m_p^\nu + h^\nu) \xi_i^\nu S^\rho] \} \gg\} \quad (\text{C.44})$$

E a função partição pode ser escrita:

$$\begin{aligned} \ll Z^n \gg = e^{-\frac{\beta pn}{2}} \int \prod_{\nu p} \frac{dm_p^\nu}{\sqrt{2\pi}} \int \prod_{(\rho\sigma)} dr_{\rho\sigma} \int \prod_{(\rho\sigma)} dq_{\rho\sigma} \exp\{N[-\frac{1}{2} \alpha \text{Tr}(\ln K) - \frac{1}{2} \alpha \beta^2 \sum_{\rho\sigma} r_{\rho\sigma} q_{\rho\sigma} - \frac{1}{2} \beta \sum_{\nu p} (m_p^\nu)^2]\} \\ \exp\{N \ll \ln \{ \text{Tr}_{S^p} \exp \beta [\frac{1}{2} \alpha \beta^2 \sum_{\rho \neq \sigma} r_{\rho\sigma} S^\rho S^\sigma + \sum_{\nu p} (m_p^\nu + h^\nu) \xi_i^\nu S^\rho] \} \gg\} \end{aligned} \quad (\text{C.45})$$

E definindo:

$$H_\xi \equiv \frac{1}{2} \alpha \beta^2 \sum_{\rho \neq \sigma} r_{\rho\sigma} S^\rho S^\sigma + \sum_{\nu p} (m_p^\nu + h^\nu) \xi_i^\nu S^\rho \quad (\text{C.46})$$

a função partição resulta:

$$\begin{aligned} \ll Z^n \gg = e^{-\frac{\beta pn}{2}} \int \prod_{\nu p} \frac{dm_p^\nu}{\sqrt{2\pi}} \int \prod_{(\rho\sigma)} dr_{\rho\sigma} \int \prod_{(\rho\sigma)} dq_{\rho\sigma} \\ \exp\{-N[\frac{1}{2} \alpha \text{Tr}(\ln K) + \frac{1}{2} \alpha \beta^2 \sum_{\rho\sigma} r_{\rho\sigma} q_{\rho\sigma} + \frac{1}{2} \beta \sum_{\nu p} (m_p^\nu)^2 - \ll \ln(\text{Tr}_{S^p} \exp(\beta H_\xi)) \gg]\} \end{aligned} \quad (\text{C.47})$$

obtemos para a equação acima:

$$\ll Z^n \gg = e^{-\frac{\beta pn}{2}} \int \prod_{\nu p} \frac{dm_p^\nu}{\sqrt{2\pi}} \int \prod_{(\rho\sigma)} dr_{\rho\sigma} \int \prod_{(\rho\sigma)} dq_{\rho\sigma} e^{-N\beta f_1(m, q, r)} \quad (\text{C.48})$$

onde

$$f_1(m, q, r) = \frac{1}{2} \sum_{\nu p} (m_p^\nu)^2 + \frac{\alpha}{2\beta} \text{Tr}(\ln K) + \frac{\alpha\beta}{2} \sum_{\rho\sigma} r_{\rho\sigma} q_{\rho\sigma} - \frac{1}{\beta} \langle\langle \ln(\text{Tr}_{S^p} \exp(\beta H_\xi)) \rangle\rangle \quad (\text{C.49})$$

O fator N no expoente permite-nos aplicar o método do ponto de sela novamente em relação as variáveis m's, q's e r's. Obtemos assim, a energia livre por unidade como:

$$f = \frac{F}{N} = -\frac{1}{\beta N} \langle\langle \ln Z \rangle\rangle = -\frac{1}{\beta N} \lim_{n \rightarrow 0} \frac{1}{n} (\langle\langle Z^n \rangle\rangle - 1) \quad (\text{C.50})$$

ou

$$f = -\frac{1}{\beta N} \lim_{n \rightarrow 0} \frac{1}{n} \ln \langle\langle Z^n \rangle\rangle = \frac{\alpha}{2} + \lim_{n \rightarrow 0} \frac{1}{n} \min f_1(m, q, r) \quad (\text{C.51})$$

Os pontos de sela podem ser determinados pelas equações:

$$\frac{\partial f}{\partial m_p^\mu} = 0 \quad \frac{\partial f}{\partial q_{\rho\sigma}} = 0 \quad \frac{\partial f}{\partial r_{\rho\sigma}} = 0 \quad (\text{C.52})$$

Como no caso mais simples de $\alpha = 0$ as equações acima conduzem a interpretação dos parâmetros de ordem $m_p^\mu, q_{\rho\sigma}, r_{\rho\sigma}$ nos pontos de sela como:

$$m_p^\mu = \frac{1}{N} \sum_i \xi_i^\mu \langle S_i^p \rangle \quad (\text{C.53})$$

Como já vimos anteriormente, o parâmetro m definido nos dá a sobreposição média dos estados visitados pela rede com os padrões armazenados.

Um segundo parâmetro de ordem é o de Edwards-Anderson (71) definido por:

$$q_{\rho\sigma} = \left\langle \left\langle \frac{1}{N} \sum_i \langle S_i^\rho \rangle \langle S_i^\sigma \rangle \right\rangle \right\rangle \quad (\text{C.54})$$

que permite discriminar uma fase paramagnética de uma fase de vidro de spin.

Por último necessitamos de outro parâmetro para descrever o ruído devido aos padrões não condensados, pois, se a rede encontra-se num estado com uma grande sobreposição com uns poucos padrões armazenados, a soma destas sobreposições aleatórias com todos os outros padrões pode gerar uma quantidade significativa de ruído.

Se s for o número de padrões condensados, isto é, com sobreposição finita à medida que N tende ao infinito, este último parâmetro é expresso por:

$$r_{p\sigma} = \frac{1}{\alpha} \sum_{\mu>s} \langle\langle m_p^\mu m_\sigma^\mu \rangle\rangle \quad (C.55)$$

Para prosseguirmos faremos a hipótese de que as réplicas do sistema são iguais, isto significa que os valores dos pontos de sela dos parâmetros de ordem não dependem dos índices das réplicas (esta hipótese não é verdadeira para temperaturas próximas de zero, mas é ainda uma boa aproximação), isto é:

$$\begin{aligned} m_p^\mu &= m^\mu \\ q_{p\sigma} &= q \\ r_{p\sigma} &= r \end{aligned} \quad (C.56)$$

Com esta simplificação os parâmetros de ordem ficam:

$$m^\mu = \langle\langle \frac{1}{N} \sum_{i=1}^N \xi_i^\mu \langle S_i \rangle \rangle\rangle \quad (C.57)$$

$$q = \langle\langle \frac{1}{N} \sum_{i=1}^N \langle S_i \rangle^2 \rangle\rangle \quad (C.58)$$

$$r = \langle\langle \frac{N}{p} \sum_{\mu=s+1}^{p-\alpha N} \langle m^\mu \rangle^2 \rangle\rangle = \frac{1}{\alpha} \sum_{\mu>s} \langle\langle [\frac{1}{N} \sum_i \xi_i^\mu \langle S_i \rangle]^2 \rangle\rangle \quad (C.59)$$

Usando então a hipótese da simetria das réplicas a equação da energia livre que é dada por:

$$f(m, q, r) = \frac{\alpha}{2} + \frac{1}{2n} \sum_{\nu p} (m_p^\nu)^2 + \frac{\alpha}{2\beta n} \text{Tr}(\ln K) + \frac{\alpha\beta}{2n} \sum_{\rho\sigma} r_{\rho\sigma} q_{\rho\sigma} - \frac{1}{\beta n} \langle\langle \ln(\text{Tr}_{\rho\sigma} \exp(\beta H_\xi)) \rangle\rangle \quad (\text{C.60})$$

fica:

$$f(m, q, r) = \frac{\alpha}{2} + \frac{1}{2} \sum_{\nu} (m^\nu)^2 + \frac{\alpha}{2\beta} \left[\ln(1 - \beta(1 - q)) - \frac{\beta q}{1 - \beta(1 - q)} \right] + \frac{\alpha\beta r}{2} (1 - q) - \beta^{-1} \int_{-\infty}^{\infty} \frac{dz}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} \langle\langle \ln 2 \cosh \beta [\sqrt{\alpha r} z + \sum_{\nu} (m^\nu + h^\nu) \xi_\nu] \rangle\rangle \quad (\text{C.61})$$

Observações para a obtenção da equação da energia livre

$$\frac{1}{2n} \sum_{\nu p} (m_p^\nu)^2 = \frac{1}{2n} \sum_{\nu} [(m_1^\nu)^2 + (m_2^\nu)^2 + \dots + (m_n^\nu)^2] = \frac{1}{2} \sum_{\nu p} (m^\nu)^2 \quad (\text{C.62})$$

$$\frac{\alpha\beta}{2n} \sum_{\rho\sigma} r_{\rho\sigma} q_{\rho\sigma} = \frac{\alpha\beta}{2n} \frac{n(n-1)}{2} r q \xrightarrow{n \rightarrow 0} -\frac{\alpha\beta}{2} r q \quad (\text{C.63})$$

$$\text{Tr}(\ln K) = \text{Tr}(\ln[(1 - \beta)I - \beta Q]) = \ln(\det[(1 - \beta)I - \beta Q]) = \ln \prod_{\rho=1}^n \lambda_\rho \quad (\text{C.64})$$

Onde λ_ρ são os autovalores da matriz K.

Não é difícil verificar que esta matriz possui um autovalor $1 - \beta - (n - 1)\beta q$ e $(n - 1)$ autovalores $1 - \beta(1 - q)$ portanto:

$$\ln \prod_{\rho=1}^n \lambda_\rho = \ln[1 - \beta - (n - 1)\beta q] + (n - 1) \ln[1 - \beta(1 - q)] \quad (\text{C.65})$$

Fazendo $a = 1 - \beta$ e $b = \beta q$ obtemos:

$$\ln \prod_{\rho=1}^n \lambda_{\rho} = \ln[a - (n-1)b] + (n-1) \ln[a+b] =$$

$$n \ln[a+b] + \ln \left[\frac{a - (n-1)b}{a+b} \right] = n \ln[a+b] + \ln \left[1 - \frac{nb}{a+b} \right]$$
(C.66)

Desenvolvendo então o logaritmo:

$$\ln \prod_{\rho=1}^n \lambda_{\rho} = n \ln[1 - \beta(1-q)] - \frac{n\beta q}{1 - \beta(1-q)}$$
(C.67)

e por fim

$$\frac{1}{\beta n} \langle \langle \ln \{ \text{Tr}_{S^{\rho}} \exp(\beta H_{\xi}) \} \rangle \rangle = \frac{1}{\beta n} \langle \langle \ln \{ \text{Tr}_{S^{\rho}} \exp \left[\frac{1}{2} \alpha \beta^2 \sum_{\rho \neq \sigma} r_{\rho\sigma} S^{\rho} S^{\sigma} + \sum_{\nu} (m_{\nu}^y + h^{\nu}) \xi_{\nu}^y S^{\rho} \right] \} \rangle \rangle$$
(C.68)

Utilizando as identidades:

$$\left(\sum_{\rho} S^{\rho} \right)^2 = \sum_{\rho=\sigma} S^{\rho} S^{\sigma} + \sum_{\rho \neq \sigma} S^{\rho} S^{\sigma} = n + \sum_{\rho \neq \sigma} S^{\rho} S^{\sigma}$$
(C.69)

$$\exp \left[\frac{1}{2} \alpha \beta^2 r \left(\sum_{\rho} S^{\rho} \right)^2 \right] = \int_{-\infty}^{\infty} \frac{dz}{\sqrt{2\pi}} \exp \left[-\frac{1}{2} z^2 + \beta \sqrt{\alpha r} z \sum_{\rho} S^{\rho} \right]$$
(C.70)

obtemos:

$$\frac{1}{\beta n} \langle \langle \ln \{ \text{Tr}_{S^{\rho}} \exp(\beta H_{\xi}) \} \rangle \rangle =$$

$$-\frac{1}{2} \alpha \beta r + \frac{1}{\beta n} \langle \langle \ln \int_{-\infty}^{\infty} \frac{dz}{\sqrt{2\pi}} e^{-\frac{1}{2} z^2} \text{Tr}_{S^{\rho}} \exp \left[\beta \left(\sqrt{\alpha r} z + \sum_{\nu} (m_{\nu}^y + h^{\nu}) \xi_{\nu}^y \sum_{\rho} S^{\rho} \right) \right] \rangle \rangle$$
(C.71)

ou ainda:

$$\begin{aligned} \frac{1}{\beta n} \ll \ln \{ \text{Tr}_{S^p} \exp(\beta H_\xi) \} \gg = \\ -\frac{1}{2} \alpha \beta r + \frac{1}{\beta n} \ll \ln \int_{-\infty}^{\infty} \frac{dz}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2} \text{Tr}_{S^p} \prod_p \exp[\beta (\sqrt{\alpha} r z + \sum_v (m^v + h^v) \xi^v S^p)] \gg \end{aligned} \quad (\text{C.72})$$

calculando o traço teremos:

$$\begin{aligned} \frac{1}{\beta n} \ll \ln \{ \text{Tr}_{S^p} \exp(\beta H_\xi) \} \gg = \\ -\frac{1}{2} \alpha \beta r + \frac{1}{\beta n} \ll \ln \int_{-\infty}^{\infty} \frac{dz}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2} [2 \cosh \beta (\sqrt{\alpha} r z + \sum_v (m^v + h^v) \xi^v)]^n \gg \end{aligned} \quad (\text{C.73})$$

e finalmente:

$$\begin{aligned} \frac{1}{\beta n} \ll \ln \{ \text{Tr}_{S^p} \exp(\beta H_\xi) \} \gg = \\ -\frac{1}{2} \alpha \beta r + \frac{1}{\beta n} \ll \ln \int_{-\infty}^{\infty} \frac{dz}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2} \exp \left[n \ln [2 \cosh \beta (\sqrt{\alpha} r z + \sum_v (m^v + h^v) \xi^v)] \right] \gg \end{aligned} \quad (\text{C.74})$$

Desenvolvendo então a exponencial obtemos:

$$\begin{aligned} \frac{1}{\beta n} \ll \ln \{ \text{Tr}_{S^p} \exp(\beta H_\xi) \} \gg = \\ -\frac{1}{2} \alpha \beta r + \frac{1}{\beta n} \ll \ln \int_{-\infty}^{\infty} \frac{dz}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2} \left[1 + n \ln [2 \cosh \beta (\sqrt{\alpha} r z + \sum_v (m^v + h^v) \xi^v)] \right] \gg \end{aligned} \quad (\text{C.75})$$

e por fim, desenvolvendo o logaritmo:

$$\begin{aligned} \frac{1}{\beta n} \ll \ln \{ \text{Tr}_{S^p} \exp(\beta H_\xi) \} \gg = \\ -\frac{1}{2} \alpha \beta r + \frac{1}{\beta} \int_{-\infty}^{\infty} \frac{dz}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2} \ll \ln [2 \cosh \beta (\sqrt{\alpha} r z + \sum_v (m^v + h^v) \xi^v)] \gg \end{aligned} \quad (\text{C.76})$$

C.3 Diagrama de fase da rede de Hopfield (106).

A partir da equação da energia livre por neurônio (C.61), do fato de que se um sistema em contato com um reservatório tem seus parâmetros externos mantidos constantes, a situação de equilíbrio estável é caracterizado pela condição de que a função energia livre de Helmholtz atinge seu valor mínimo (100-102), isto é:

$$\Delta F_0 \leq 0 \quad (C.77)$$

das relações

$$\frac{\partial f}{\partial m_p^\mu} = 0 \quad \frac{\partial f}{\partial q_{\rho\sigma}} = 0 \quad \frac{\partial f}{\partial r_{\rho\sigma}} = 0 \quad (C.78)$$

e supondo serem as réplicas simétricas, obtemos para os parâmetros de ordem:

$$m^v = \langle\langle \xi^v \operatorname{tgh} \beta [\sqrt{\alpha} r z + (\bar{m} + \bar{h}) \cdot \xi^-] \rangle\rangle \quad (C.79)$$

$$q = \langle\langle \operatorname{tgh}^2 \beta [\sqrt{\alpha} r z + (\bar{m} + \bar{h}) \cdot \xi^-] \rangle\rangle \quad (C.80)$$

$$r = \frac{q}{[1 - \beta(1 - q)]^2} \quad (C.81)$$

Observações:

O cálculo da derivada da função energia livre por neurônio em relação ao parâmetro r conduz-nos:

$$\frac{1}{2} \alpha \beta^2 (q - 1) = -\frac{\beta \sqrt{\alpha}}{2\sqrt{r}} \int_{-\infty}^{\infty} \frac{dz}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} z \langle\langle \xi^v \operatorname{tgh} \beta [\sqrt{\alpha} r z + (\bar{m} + \bar{h}) \cdot \xi^-] \rangle\rangle_\xi \quad (C.82)$$

onde as duas médias foram separadas temporariamente. Integrando por partes e utilizando a identidade:

$$\int_{-\infty}^{\infty} \frac{dz}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} z f(z) = \int_{-\infty}^{\infty} \frac{dz}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} f'(z) \quad (C.83)$$

obtemos a equação para o parâmetro q .

Finalmente derivando f em relação ao parâmetro q obtemos a equação para r .

A solução numérica e analítica das equações acima (C.79-C.81) produz o diagrama de fase da rede de Hopfield como na figura a seguir.

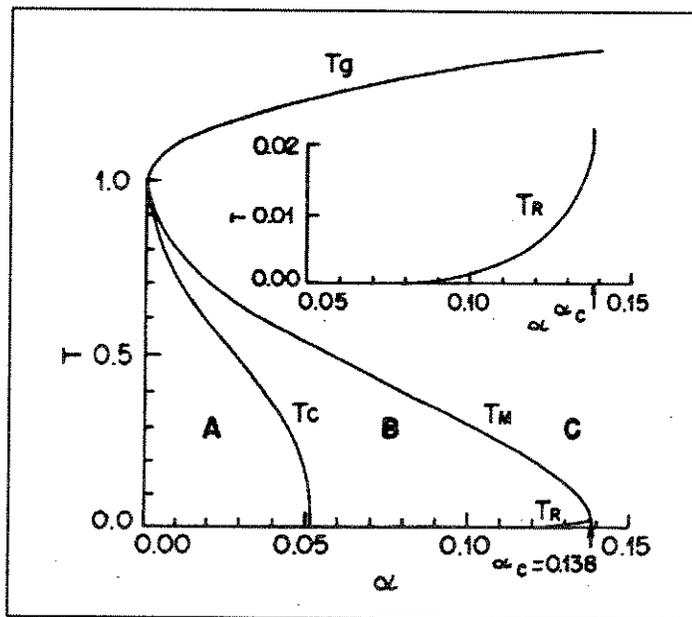


Figura C.1 - Diagrama de Fase do Modelo de Hopfield - (Ref. (70))

com as seguintes observações:

1. Para $\alpha < 0,138$ (fase ferromagnética) existe a recuperação dos padrões armazenados. (**Região A+B**)

2. Para $\alpha > 0,138$ (fase de vidro de spin) não há recuperação dos padrões armazenados. (**Região C**)

3. Para $\alpha < 0,051$ os padrões armazenados são mínimos globais. (**Região A**)

4. Para $0,051 < \alpha < 0,138$ os padrões armazenados tornam-se mínimos locais e coexistem com outro tipo de atrator denominado do tipo vidro de spin que possuem energia menor que a dos padrões armazenados. (**Região B**)

5. Numa tentativa de visualização da topologia da função energia livre nas diferentes partes do diagrama de fase obtemos:

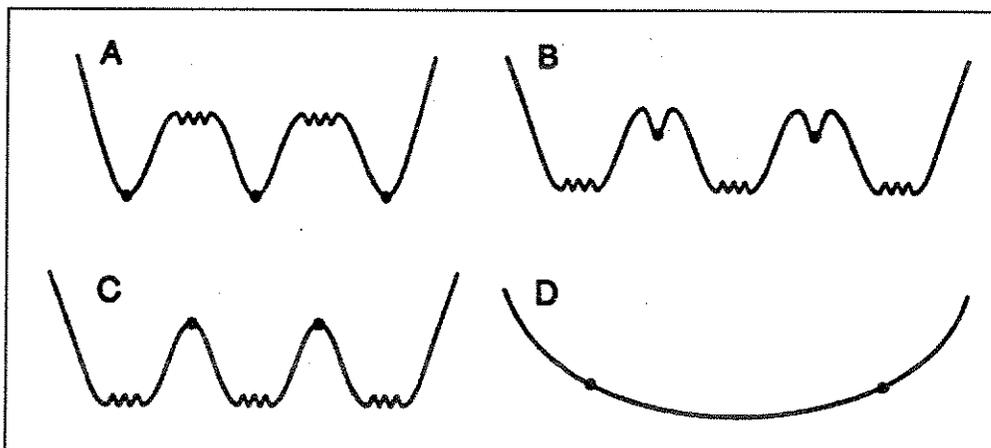


Figura C.2 - Representação Unidimensional da Energia Livre nas Várias Regiões do Diagrama de Fase - (Ref. (111))

6. Ao passarmos da região C para B, o parâmetro m muda descontinuamente para o valor zero (transição de primeira ordem) exceto para $\alpha=0$ e $T=1$ onde a transição é de segunda ordem.

APENDICE D.

APLICAÇÃO DOS MÉTODOS DA MECÂNICA ESTATÍSTICA AO MODELO PROPOSTO.

Neste modelo consideramos uma rede na qual a interação entre os neurônios é descrita por:

$$J_{ij} = \frac{1}{N} \sum_{\mu=1}^P \xi_i^{\mu} \xi_j^{\mu} [\cos k(x_i - x_j)] \quad (D.1)$$

Vamos estudar o modelo, supondo que não haja saturação ($\alpha = 0$). Um estudo considerando a saturação do modelo, poderá ser feito posteriormente.

Logo para o Hamiltoniano teremos:

$$H = -\frac{1}{2N} \sum_{ij} \left[\sum_{\mu=1}^P \xi_i^{\mu} \xi_j^{\mu} \cos k(x_i - x_j) \right] S_i S_j \quad (D.2)$$

Usando a identidade:

$$\cos k(x_i - x_j) = \cos kx_i \cos kx_j + \sin kx_i \sin kx_j \quad (D.3)$$

e definindo:

$$\begin{aligned} u_i^{(1\mu)} &= \xi_i^\mu \cos kx_i & \phi_i^{(1\mu)} &= u_i^{(1\mu)} S_i \\ u_i^{(2\mu)} &= \xi_i^\mu \sin kx_i & \phi_i^{(2\mu)} &= u_i^{(2\mu)} S_i \end{aligned} \quad (D.4)$$

obtemos:

$$H = -\frac{1}{2N} \sum_{n=1}^2 \sum_{\mu=1}^p [\sum_{ij} u_i^{(n\mu)} u_j^{(n\mu)}] S_i S_j \quad (D.5)$$

Usando a identidade:

$$[\sum_{i=1}^N u_i^{(n\mu)} S_i]^2 = \sum_{ij} u_i^{(n\mu)} u_j^{(n\mu)} S_i S_j + \sum_{i=1}^N (u_i^{(n\mu)} S_i)^2 \quad (D.6)$$

então

$$H = -\frac{1}{2N} \sum_n \sum_\mu [\sum_{i=1}^N u_i^{(n\mu)} S_i]^2 - \frac{1}{2N} \sum_n \sum_\mu \sum_{i=1}^N (u_i^{(n\mu)} S_i)^2 \quad (D.7)$$

Efetuada a somatória sobre o segundo membro da equação (D.7) acima temos:

$$H = -\frac{1}{2N} \sum_n \sum_\mu [\sum_{i=1}^N u_i^{(n\mu)} S_i]^2 - \frac{1}{2} p \quad (D.8)$$

ou ainda

$$H = -\frac{1}{2N} \sum_n \sum_\mu [\sum_i \phi_i^{(n\mu)}]^2 - \frac{1}{2} p \quad (D.9)$$

A função partição será expressa por:

$$Z = \text{Tr}_S \exp\left[-\frac{\beta}{2N} \sum_n \sum_\mu \left[\sum_i \phi_i^{(n\mu)}\right]^2\right]$$

$$Z = \text{Tr}_S \exp\left[\sum_n \sum_\mu \left[\sum_i \left(\frac{\beta}{2N}\right)^{\frac{1}{2}} \phi_i^{(n\mu)}\right]^2\right] \quad (\text{D.10})$$

Empregando as propriedades das integrais gaussianas(107,108)

$$\exp(\lambda a^2) = \left(\frac{1}{2\pi}\right)^{\frac{1}{2}} \int_{-\infty}^{\infty} dx \exp\left[-\frac{x^2}{2} + \sqrt{2\lambda} ax\right] \quad (\text{D.11})$$

obtemos:

$$Z = \text{Tr}_S \int \prod_{n\mu} \frac{dx^{(n\mu)}}{\sqrt{2\pi}} \exp\left[-\frac{1}{2} \sum_n \sum_\mu (x^{(n\mu)})^2 + \sum_n \sum_\mu \sum_i \left(\frac{\beta}{N}\right)^{\frac{1}{2}} \phi_i^{(n\mu)} x^{(n\mu)}\right] \quad (\text{D.12})$$

Definindo as variáveis

$$x^{(n\mu)} = \sqrt{N} y^{(n\mu)} \quad (\text{D.13})$$

a função partição fica:

$$Z = \text{Tr}_S \int \prod_{n\mu} \sqrt{\frac{N}{2\pi}} dy^{(n\mu)} \exp\left[-N \sum_n \sum_\mu \frac{(y^{(n\mu)})^2}{2} + \sum_n \sum_\mu \sum_i \beta^{\frac{1}{2}} \phi_i^{(n\mu)} y^{(n\mu)}\right] \quad (\text{D.14})$$

Usando a identidade

$$\text{Tr}_S \exp\left[\sum_i f(S_i)\right] = \exp\left[\ln(\text{Tr}_S \exp[\sum_i f(S_i)])\right] \quad (\text{D.15})$$

e definindo:

$$G = \sum_n \sum_\mu \frac{(y^{(n\mu)})^2}{2} - \frac{1}{N} \ln(\text{Tr}_S \exp[\sum_n \sum_\mu \sum_i \beta^{\frac{1}{2}} \phi_i^{(n\mu)} y^{(n\mu)}]) \quad (\text{D.16})$$

temos para a função partição:

$$Z = \int \prod_{n\mu} \sqrt{\frac{N}{2\pi}} dy^{(n\mu)} \exp[-NG(y^{(n\mu)})] \quad (\text{D.17})$$

Definindo

$$A_i(S_i) = \exp[\sum_n \sum_\mu \beta^{\frac{1}{2}} \phi_i^{(n\mu)} y^{(n\mu)}] \quad (\text{D.18})$$

aplicando o método do "steepest descent" (109,110)

$$\frac{\partial G}{\partial y^{(n\mu)}} = y^{(n\mu)} - \frac{1}{N} \sum_i \beta^{\frac{1}{2}} \frac{\text{Tr}_S[\phi_i^{(n\mu)} A_i(S_i)]}{\text{Tr}_S[A_i(S_i)]} = 0 \quad (\text{D.19})$$

e definindo as variáveis

$$y^{(n\mu)} = \beta^{\frac{1}{2}} m^{(n\mu)} \quad (\text{D.20})$$

$$B_i(S_i) = \exp[\sum_n \sum_\mu \beta \phi_i^{(n\mu)} m^{(n\mu)}] \quad (\text{D.21})$$

obtemos

$$m^{(n\mu)} = \frac{1}{N} \sum_i \frac{\text{Tr}_S[\phi_i^{(n\mu)} B_i(S_i)]}{\text{Tr}_S[B_i(S_i)]} \quad (\text{D.22})$$

Ao efetuarmos o traço chegamos ao conjunto de equações:

$$m^{(1\mu)} = \frac{1}{N} \sum_i \xi_i^\mu \cos kx_i \text{tgh } \beta \left[\sum_\mu \xi_i^\mu (m^{(1\mu)} \cos kx_i + m^{(2\mu)} \text{sen } kx_i) \right] \quad (\text{D.23})$$

$$m^{(2\mu)} = \frac{1}{N} \sum_i \xi_i^\mu \operatorname{sen} kx_i \operatorname{tgh} \beta \left[\sum_\mu \xi_i^\mu (m^{(1\mu)} \cos kx_i + m^{(2\mu)} \operatorname{sen} kx_i) \right]$$

(D.24)

BIBLIOGRAFIA

(1) M. Fibich, I. Ries e A. Aron.

Thermodynamics properties of spin glasses.

Journal of Magnetism and Magnetic Materials.7,262-264,1978.

(2) J. D. Cowan, D. H. Sharp

Neural nets.

Quartely Reviews Of Biophysics, 21(3),365-427,1988.

(3) S. R. Cajal.

Histology du Système Nerveux de l'Homme et des Vertébrés.

Consejo Superior de Investigaciones Científicas.

Instituto Ramon y Cajal, Madrid, 1952

(4) E. D. Adrian.

The all-or-none principle in nerve.

Journal Physiology, Lond. XLVII,(6),460-474,1914.

(5) M. Santini (ed.)

Golgi Centennial Symposium.

Raven Press, New York, 1975.

(6) J. C. Eccles.

The Physiology of Nerve Cells.

John Hopkins University Press, 1957.

(7) J. C. Eccles, M. Ito, J. Szentagothai.

The Cerebellum as a Neuronal Machine.

Springer-Verlag, New York, 1967.

-
- (8) D. A. Scholl.
The Organization of the Cerebral Cortex.
Methuen, London, 1956.
- (9) J. Szentágothai, M.A. Arbib.
Conceptual models of neural organization.
Neurosciences Research Program Bulletin, 12,3,1974.
- (10) J. T. Schmidt (ed.)
Activity-dependent synaptic changes.
Cellular and Molecular Biology, 5,1-2,1985.
- (11) R. M. Gaze.
The Formation of Nerve Connections.
Academic Press, New York, 1970.
- (12) A. L. Hodgkin, A.F. Huxley.
A quantitative description of membrane current and its application to conduction and excitation in nerve.
Journal of Physiology, Lond. 117,500-544,1952.
- (13) T. J. Sejnowski, C. Koch, P. S. Churchland.
Computational neuroscience.
Science, 241,1299-1306,1988.
- (14) C. C. Klimasauskas, J. P. Guiver.
NeuralWorks.
NeuralWare Inc, 1988.
- (15) W. S. McCulloch, W. Pitts.
A logical calculus of the ideas immanent in nervous activity.
Bulletin of Mathematical Biophysics, 5,115-133,1943.

-
- (16) W. Pitts, W. S. McCulloch.
How we know universals: the perception of auditory and visual forms.
Bulletin of Mathematical Biophysics, 9,127-147,1947.
- (17) D. O. Hebb.
The Organization of Behavior.
Wiley, New York, 1949.
- (18) K. H. Lashley.
In search of the engram.
Society of Experimental Biology Symposium, No. 4:
Psychological Mechanisms in Animal Behavior.
Cambridge University Press, 1950.
- (19) H. Barlow.
Single units and sensation: a neuron doctrine for perceptual psychology.
Perception, 1,371-394,1971.
- (20) H. Barlow.
The role of single neurons in the psychology of perception.
Quarterly Journal of Experimental Psychology, A37,121-145,1985
- (21) B. G. Cragg, H. N. V. Temperley.
The organization of neurons: a cooperative analogy.
Electroencephalogram and Clinical Neurophysiology, 6,85-92,1954
- (22) B. G. Cragg, H. N. V. Temperley.
Memory: the analogy with ferromagnetic hysteresis.
Brain, 78,304-316,1955.
- (23) B. Farley, W. A. Clark.
Simulation of self-organizing systems by digital computer.
IRE Transactions on Information Theory, 4,76-84,1954.

(24) A. M. Uttley.

The Classification of signals in the nervous system.

Electroencephalogram and Clinical Neurophysiology, 6,479-494,1954

(25) N. Rochester, J. H. Holland, L. H. Haibt, e W. L. Duda.

Tests on a cell assembly theory of the action of the brain, using a large digital computer.

IRE Transactions on Information Theory IT-2,80-93,1956.

(26) F. Rosenblatt.

The perceptron: a probabilistic model for information storage and organization in the brain.

Psychological Review, 65,386-408,1958.

(27) M. Minsky, S. Papert.

Perceptrons.

MIT Press, Cambridge, 1969.

(28) H. D. Block

The perceptron: a model for brain functioning. I.

Reviews of Modern Physics, 34(1),123-135,1962.

(29) H. D. Block, B. W. Knight, F. Rosenblatt.

Analysis of a four-layer series-coupled perceptron. II.

Reviews of Modern Physics, 34(1),135-142,1962.

(30) O. G. Selfridge.

Pandemonium: a paradigm for learning.

Mechanisation of Thought Processes:

Proceedings of a Symposium Held at the National Physical Laboratory.

London, 513-526,1958.

-
- (31) B Widrow, M. E. Hoff.
Adaptive switching circuits.
IRE WESCON Convention Record, NY, IRE,96-104,1960.
- (32) B. Widrow, R. Winter.
Neural nets for adaptive filtering and adaptive pattern recognition.
Computer, 25-39,Mar/1988.
- (33) B. Widrow, M. A. Lehr.
30 years of adaptive neural networks: perceptron, madaline, and backpropagation.
Proceedings of the IEEE, 78(9),1415-1442,1990.
- (34) B. Widrow, S. D. Stearns.
Adaptive Signal Processing.
Prentice-Hall, New York,1985.
- (35) D. E. Rumelhart, G. E. Hinton, R. J. Williams.
Learning internal representation by error propagation.
Parallel Distributed Processing.
D. E. Rumelhart, J. L. McClelland (ed.),1986.
- (36) T. Kohonen.
Correlation matrix memories.
IEEE Transactions on Computers, C-21,353-359,1972.
- (37) J. A. Anderson.
A simple neural network generating an interactive memory.
Mathematical Biosciences, 14,197-220,1972.
- (38) W. A. Little.
The existence of persistent states in the brain.
Mathematical Biosciences, 19,101-120,1974.

(39) W. A. Little, G. L. Shaw.

A statistical theory of short and long term memory.

Behavioral Biology, 14,115-133,1975

(40) D. H. Acley, G. E. Hinton, T. J. Sejnowski.

A learning algorithm for Boltzmann machines.

Cognitive Science, 9,147-169,1985.

(41) D. Marr, T. Poggio.

Cooperative computation of stereo disparity.

Science, 194,283-287,1976.

(42) J. J. Hopfield, D. W. Tank.

"Neural" computation of decisions in optimization problems.

Biological Cybernetics, 52,141-152,1985.

(43) S. -I. Amari.

Neural theory of association and concept formation.

Biological Cybernetics, 26,175-185,1977.

(44) J. A. Anderson, J. W. Silverstein, S. A. Ritz, R. S. Jones.

Distinctive features, categorical perception, and probability learning: some applications of a neural model

Psychological Review, 84(5),413-451,1977.

(45) L. N. Cooper.

A possible organization of animal memory and learning.

Proceedings of the Nobel Symposium on Collective Properties of Physical Systems.

B. Lundquist, S. Lundquist, (eds.).

Academic Press, 252-264,1973.

(46) M. M. Nass, L. N. Cooper.

A theory for the development of feature detecting cells in visual cortex.
Biological Cybernetics, 19,1-18,1975.

(47) F. Ratliff, B. W. Knight, F. A. Dodge, H. K. Hartline.

Fourier analysis of dynamics of excitation and inhibition in the eye of Limulus: amplitude, phase, and distance.
Vision Research, 14,1155-1168,1974.

(48) G. Stephenson.

Introdução a Matrizes, Conjuntos e Grupos.
Edgard Blücher Ltda, 1975.

(49) J. J. Hopfield.

Neural networks and physical systems with emergent collective computational abilities.
Proceedings of the National Academy of Sciences, 79,2554-2558,1982.

(50) J. F. Fontanari.

Processamento de informações em redes de neurônios síncronas.
Tese de doutorado, USP-IFQ-São Carlos,1988.

(51) G. Toulouse, S. Dehaene, J. -P. Changeaux.

Spin glass model of learning by selection.
Proceedings of the National Academy of Science USA,83,1695-1698,1986.

(52) M. A. Cohen, S. Grossberg.

Absolute stability of global pattern formation and parallel memory storage by competitive neural networks.
IEEE Transactions on Systems, Man, and Cybernetics, 13,815-826,1983.

(53) R. J. Glauber.

Time-dependent statistics of the Ising model.
Journal of Mathematical Physics, 4,294-307,1963

(54) H. E. Stanley.

Phase Transitions and Critical Phenomena.

Oxford University Press, 1971.

(55) D. C. Mattis.

Solvable spin systems with random interactions.

Physics Letters, 56A(5),421-422,1976.

(56) G. E. Hinton, T. J. Sejnowski.

Optimal perceptual inference.

Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition.

Washington 1983,448-453,

(57) P. Peretto.

Collective properties of neural networks: a statistical physics approach.

Biological Cybernetics, 50,51-62,1984.

(58) F. Reif

Fundamentals of Statistical and Thermal Physics.

McGraw-Hill Kogakusha, LTD, 1965.

(59) H. Cramer.

Metodos Matematicos de Estadistica.

Aguilar S. A. de Ediciones, 1953.

(60) P. A. Pury.

Asymmetry and convergence in the central limit theorem: an approach for physicist.

American Journal of Physics, 58(1),62-67,1990.

(61) H. J. Larson, B. O. Shubert.

Probabilistic Models in Engineering Sciences.

John Wiley & Sons, 1979.

-
- (62) V. P. Spiridonov, A. A. Lopatkin.
Tratamiento Matemático de Datos Físico-Químicos.
Editorial Mir, Moscú, 1973.
- (63) A. Papoulis.
The Fourier Integral and its Applications.
McGraw-Hill Book Company, 1962.
- (64) T. N. Thiele.
Theory of Observations.
Londres, 1903.
- (65) R. Kubo.
Generalized cumulant expansion method.
Journal of the Physical Society of Japan, 17(7),1100-1120,1962.
- (66) G. Horwitz, H. B. Callen.
Diagrammatic expansion for the Ising model with arbitrary spin and range of interaction.
Physical Review, 124(6),1757-1785,1961.
- (67) G. S. Rushbrooke.
On the theory of randomly dilute Ising and Heisenberg ferromagnets.
Journal of Mathematical Physics, 5(8),1106-1116,1964.
- (68) S. A. Rice.
The Statistical Mechanics of Simple Liquids.
Interscience Publishers, 1965.
- (69) E. Meeron.
Series expansion of distribution functions in multicomponent fluid systems.
Journal of Chemical Physics, 27(6),1238-1246,1957.

(70) D. J. Amit.

Modeling Brain Function.

Cambridge University Press, 1989.

(71) S. F. Edwards, P. W. Anderson.

Theory of spin glasses.

Journal of Physics F (Metal Physics), 5, 965-974, 1975

(72) P. Peretto.

On learning rules and memory storage abilities of asymmetrical neural networks.

Journal de Physique France, 49, 711-726, 1988.

(73) E. Domany, W. Kinzel, R. Meir.

Layered neural networks.

Journal Physics A (Mathematics and General), 22, 2081-2102, 1989.

(74) D. J. Amit, H. Gutfreund, H. Sompolinsky.

Statistical mechanics of neural networks near saturation.

Annals of Physics, 173, 30-67, 1987.

(75) J. P. Nadal, G. Toulouse, J. P. Changeaux, S. Dehaene.

Networks of formal neurons and memory palimpsests.

Europhysics Letters, 1, 535-542, 1986.

(76) D. J. Amit, H. Gutfreund, H. Sompolinsky.

Information storage in neural networks with low levels of activity.

Physical Review, 35A(5), 2293-2303, 1987.

(77) J. F. Fernández, D. Sherrington.

Randomly located spins with oscillatory interactions.

Physical Review 18B(11), 6270-6274, 1978.

-
- (78) H. Orland, C. De Dominicis, T. Garel.
Non-homogenous mean field picture for spin-glasses.
Journal de Physique - Lettres, 42,L73-L77,1981.
- (79) M. V. Feigel'man, L. B. Ioffe.
One-dimensional spin glass with oscillating long-range interaction.
Zeitschrift für Physik B - Condensed Matter, 51,237-249,1983.
- (80) K. H. Fisher, J.A. Hertz.
Spin Glasses.
Cambridge University Press, 1991
- (81) D. Chowdhury.
Spin Glasses and Other Frustrated Systems.
Princeton University Press, 1986.
- (82) K. Binder, A. P. Young.
Spin glasses: experimental facts, theoretical concepts.
Reviews of Modern Physics, 58(04),801-876,1986
- (83) M. Mezard, G. Parisi, M. A. Virasoro.
Spin Glass Theory and Beyond.
World Scientific Publishing Co., Inc, 1987
- (84) J. L. van Hemmen, I. Morgenstern (ed.).
Heidelberg Colloquium on Spin Glasses.
Lectures Notes in Physics n.192
Springer-Verlag, 1983
- (85) D. Chowdhury, A. Mookerjee.
Mean-field theory of spin glasses.
Physics Reports, 114(01),1-98,1984

(86) G. Gresting, R. Kühn.

On classical spin-glass models.

Journal de Physique, 48,713-721,1987.

(87) K. Binder.

Statistical mechanics of Ising spin glasses.

Fundamental Problems in Statistical Physics.

E. G. D. Cohen (ed.),21-51,1980.

(88) J. L. van Hemmen.

Classical spin-glass model.

Physical Review Letters, 49(6),409-412,1982.

(89) W. Kinzel.

Spin glasses and memory.

Physica Scripta, 35,398-401,1987.

(90) J. P. Provost, G. Vallée.

*Ergodicity of the coupling constants and the symmetric n-replicas trick
for a class of mean-field spin-glass models.*

Physical Review Letters, 50(8),598-600,1983.

(91) A. S. Chakravarty.

Introduction to the Magnetic Properties of Solids.

John Wiley & Sons, 1980

(92) B. R. Coles.

An historical introduction on spin glasses.

Annales de Physique (France), 10,63-67,1985

(93) P W. Anderson.

Spin glass III: theory raises its head.

Physics Today, Jun/1988,9.

(94) V. Cannella, J. A. Mydosh e J. I. Budnick.

Transition metal alloys.

Journal of Applied Physics, 42(4),1689-1690,1971.

(95) M. A. Ruderman, C. Kittel.

Indirect exchange coupling of nuclear magnetic moments by conduction electrons.

Physical Review, 96(1),99-102,1954.

(96) T. Katsuya.

A theory of metallic ferro- and antiferromagnetism on Zener's model.

Progress in Theoretical Physics,16(1),45-57,1956.

(97) K. Yosida.

Magnetic properties of Cu-Mn alloys.

Physics Review,106(5),893-898,1957.

(98) G. Toulouse.

Theory of the frustration effect in spin glasses: I.

Communications in Physics,2(4),115-119,1977.

(99) S. Kirkpatrick, D. Sherrington.

Infinite-ranged models of spin-glasses.

Physical Review B.17(11),4384-4403,1978.

(100) R. P. Feynmann.

Statistical Mechanics.

W. A. Benjamin, Inc., 1972.

(101) R. Abe, Y. Takahashi.

Statistical Mechanics.

University of Tokio Press, 1975.

(102) R. K. Pathria.

Statistical Mechanics.

Pergamon Press, 1972

(103) T. F. Lin.

Problem of the disordered chain.

Journal of Mathematical Physics, 11, 1584-1590, 1970.

(104) G. Grinstein, A. H. Luther.

Application of the renormalization group to the phase transitions in disordered systems.

Physical Review B13(3), 1329-1343, 1976.

(105) V. J. Emery.

Critical properties of many-component systems.

Physical Review B11(1), 239-247, 1975

(106) D. J. Amit, H. Gutfreund, H. Sompolinsky.

Spin-glass models of neural networks.

Physical Review, 32A(2), 1007-1018, 1985.

(107) J. Hubbard.

Calculation of partition functions.

Physical Review Letters, 3(2), 77-78, 1959.

(108) R. L. Stratnovich.

On a method of calculating quantum distribution functions.

Soviet Physics Doklady, 2, 416-419, 1957.

(109) I. V. Lindell.

Delta function expansions, complex delta functions and the steepest descent method.

American Journal of Physics, 61(5), 438-442, 1993.

(110) K. Huang.

Statistical Mechanics.

John Wiley & Sons, 1963.

(111) J. Hertz, A. Krogh, R. G. Palmer.

Introduction to the Theory of Neural Computation.

Addison-Wesley Publishing Company, 1991.

(112) T. Gezti.

Physical Models of Neural Networks.

World Scientific Publishing Co., 1990

(113) Y. Kamp, M. Hasler.

Recursive Neural Networks for Associative Memory.

John Wiley & Sons, 1990.

(114) D. J. Amit.

The properties of models of simple neural networks.

Heidelberg Colloquium in Glassy Dynamics.

Lectures Notes in Physics 275,430-484,1986.

(115) L. Garrido (ed.).

Statistical Mechanics of Neural Networks.

Lectures Notes in Physics, 368.

Springer-Verlag, 1990.

(116) J. A. Hertz.

Statistical mechanics of neural computation.

The International Journal of Supercomputer Applications, 2(4),54-62,1988.

(117) W. Kinzel.

Statistical mechanics of neural networks.

Physica Scripta, T25,144-148,1989.

(118) R. Köberle.

Neural networks as content addressable memories and learning machines.
Computer Physics Communications, 56,43-50,1989.

(119) A. Zippelius.

Statistical mechanics of neural networks.
Physica 194A,471-481,1993.

(120) H. Gutfreund.

From statistical mechanics to neural networks and back.
Physica 163A,373-385,1990.

(121) E. Domany, J. L. van Hemmen, K. Schulten (eds.).

Models of neural networks.
Springer-Verlag, 1991.

(122) W. Kinzel.

Spin glasses as model systems for neural networks.
Complex Systems - Operational Approaches.
H. Haken (ed.), 107-115,1985.

(123) H. Sompolinsky.

Statistical mechanics of neural networks.
Physics Today,70-80,Dec/1988.

(124) E. Domany.

Neural networks: a biased overview.
Journal of Statistical Physics, 51(5-6),743-775,1988.

(125) Y. Fu, P.W. Anderson.

Application of statistical mechanics to NP-complete problems in combinatorial optimization.
Journal Physics A (Mathematics and General), 19,1605-1620,1986.

-
- (126) K. Y. M. Wong, D. Sherrington.
Graph bipartitioning and spin glasses on a random network of fixed finite valence.
Journal Physics A (Mathematics and General), 20, L793-L799, 1987.
- (127) K. Y. M. Wong, D. Sherrington, P. Mottishaw, R. Dewar, C. De Dominicis.
Graph partitioning and dilute spin glasses : the minimum cost solution.
Journal Physics A (Mathematics and General), L99-L105, 1988.
- (128) Y. Y. Goldschmidt, C. De Dominicis.
Replica symmetry breaking in the spin-glass model on lattices with finite connectivity: applications to graph partitioning.
Physical Review, 41B(4), 2184-2197, 1990.
- (129) G. L. Bilbro, W. E. Snyder, S. J. Garnier, J. W. Gault.
Mean field annealing: a formalism for construting GNC-like algorithms.
IEEE Transactions on Neural Networks, 3(1), 131-138, 1992
- (130) D. S. Johnson, C. A. Aragon, L. A. McGeoch, C. Schevon.
Optimization by simulated annealing: an experimental evaluation; Part 1, graph partitioning.
Operations Research, 37(6), 865-892, 1989.
- (131) L. Personnaz, I. Guyon, G. Dreyfus.
Collective computational properties of neural networks: new learning mechanisms.
Physical Review 34A(5), 4217-4228, 1986.
- (132) J. F. Fontanari.
Generalization in a Hopfield network.
Journal de Physique, 51, 2421-2430, 1990.

(133) I. Kanter, H. Sompolinsky.

Associative recall of memory without errors.

Physical Review, 35A(1),380-392,1987.

(134) H. Gutfreund.

Neural networks with hierarchically correlated patterns.

Physical Review, 37A(2),570-577,1988.

(135) R. Kühn.

Statistical mechanics for networks of analog neurons.

Lectures Notes in Physics, 368.

Statistical Mechanics of Neural Networks.

L. Garrido (ed.)

Springer-Verlag, 1990.

(136) J. L. van Hemmen, R. G. Palmer.

What's wrong with the replica method in spin glass theory ?

Colloquia Mathematica Societatis Janos Bolyai.

Random Fields, Esztergom (Hungary), 1979.

(137) J. L. van Hemmen, R. G. Palmer.

The thermodynamic limit and the replica method for short-range random systems.

Journal of Physics A (Mathematical and General), 15,3881-3890,1982.

(138) J. L. van Hemmen, A. Süto.

Quenched and annealed free energies.

Journal de Physique, 45,1277-1281,1984.

(139) N. Sourlas.

Statistical mechanics and the travelling salesman problem.

Europhysics Letters, 2(12),912-923,1986

-
- (140) M. Mezard, G. Parisi.
A replica analysis of the travelling salesman problem.
Journal de Physique, 47,1285-1296,1986.
- (141) J. Vannimenus, M. Mezard.
On the statistical mechanics of optimization problems of the travelling salesman type.
Journal de Physique-Lettres, 45,L1145-L1153,1984.
- (142) W. Krauth, M. Mezard.
The cavity method and the travelling-salesman problem.
Europhysics Letters, 8(3),213-218,1989.
- (143) E. Bonomi, J. -L. Lutton.
The N-city travelling salesman problem: statistical mechanics and the Metropolis algorithm.
SIAM Review, 26(4),551-568,1984.
- (144) A. Crisanti, G. Paladin, H. -J. Sommers, A. Vulpiani.
Replica trick and fluctuations in disordered systems.
Journal de Physique I France, 2,1325-1332,1992.
- (145) H. A. Gersch.
Cumulant expansions in statistical and quantum physics.
American Journal of Physics, 59(3),273-277,1991.
- (146) D. Bormann.
Dynamical equivalence in a class of nonlinear neural network models.
Zeitschrift für Physik B- Condensed Matter, 79,307-315,1990.
- (147) T. Kohonen.
Self-organization and associative memory.
Springer Verlag, Berlin,1984

(148) J. R. L. Almeida.

On the 1D Ising spin glass system with random long-ranged interactions.

Journal of Physics A (Mathematical and General), 24,L685-L689,1991.