



**UNIVERSIDADE FEDERAL DO PARÁ**

**CENTRO DE GEOCIÊNCIAS**

**CURSO DE PÓS-GRADUAÇÃO EM GEOFÍSICA**

**DISSERTAÇÃO DE MESTRADO**

**IDENTIFICAÇÃO AUTOMÁTICA DAS PRIMEIRAS  
QUEBRAS EM TRAÇOS SÍSMICOS POR MEIO DE UMA  
REDE NEURAL DIRETA.**

**ANNA ILCÉA FISCHETTI MIRANDA**

**BELÉM  
2000**



UNIVERSIDADE FEDERAL DO PARÁ  
CENTRO DE GEOCIÊNCIAS  
CURSO DE PÓS-GRADUAÇÃO EM GEOFÍSICA

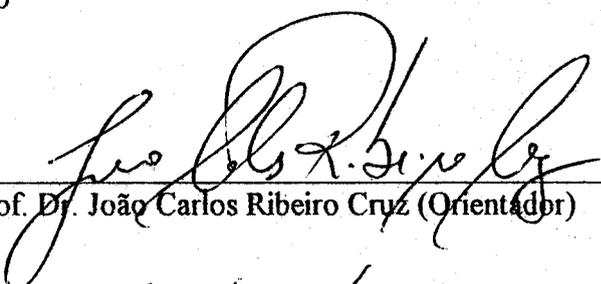
**IDENTIFICAÇÃO AUTOMÁTICA DAS PRIMEIRAS  
QUEBRAS EM TRAÇOS SÍSMICOS POR MEIO DE UMA  
REDE NEURAL DIRETA.**

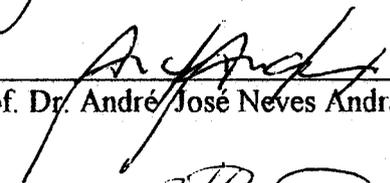
DISSERTAÇÃO APRESENTADA POR  
**ANNA ILCÉA FISCHETTI MIRANDA**

COMO REQUISITO PARCIAL À OBTENÇÃO DE GRAU DE MESTRE EM CIÊNCIAS NA  
ÁREA DE GEOFÍSICA

Data de aprovação: / /2000

COMITÊ DE TESE:

  
\_\_\_\_\_  
Prof. Dr. João Carlos Ribeiro Cruz (Orientador)

  
\_\_\_\_\_  
Prof. Dr. André José Neves Andrade (Co-orientador)

  
\_\_\_\_\_  
Prof. Dr. Ubiratan Holanda Bezerra

**BELÉM  
2000**

## **AGRADECIMENTOS**

**Ao DGf/CG/UFPA e CNPQ pelo suporte dado à pesquisa.**

# Sumário

<b>AGRADECIMENTOS</b> .....	i
<b>LISTA DE ILUSTRAÇÕES</b> .....	iv
<b>RESUMO</b> .....	1
<b>ABSTRACT</b> .....	2
<b>1- INTRODUÇÃO</b> .....	3
<b>2- MÉTODOS SÍSMICOS</b> .....	6
2.1- INTRODUÇÃO.....	6
2.2- ONDAS ACÚSTICAS.....	6
2.3- O MÉTODO SÍSMICO DE REFLEXÃO.....	8
2.4- O MÉTODO SÍSMICO DE REFRAÇÃO.....	12
2.5- SÍSMICA DE POÇO.....	15
<b>3- REDES NEURAIS ARTIFICIAIS</b> .....	20
3.1- INTRODUÇÃO.....	20
3.2- O NEURÔNIO BIOLÓGICO.....	24
3.3- O NEURÔNIO ARTIFICIAL.....	27
3.3.1- <b>Funções de ativação</b> .....	29
3.4- REDES NEURAIS ARTIFICIAIS.....	32
3.4.1- <b>Treinamento ou aprendizagem</b> .....	35
3.4.2- <b>Operação</b> .....	37
3.4.3- <b>Classificação geral das redes neurais</b> .....	38
3.4.4- <b>Generalização</b> .....	39
3.5- REDE NEURAL DIRETA MULTICAMADAS.....	40
3.5.1- <b>Aprendizagem por retro-propagação do erro</b> .....	45
3.5.1.1- <b>Algoritmo de retro-propagação do erro</b> .....	47
<b>4- IDENTIFICAÇÃO DAS PRIMEIRAS QUEBRAS COM A REDE NEURAL MULTICAMADAS</b> .....	53
4.1- INTRODUÇÃO.....	53
4.2- PRIMEIRAS QUEBRAS.....	53
4.3- APROXIMAÇÃO DE FUNÇÕES.....	55
4.3.1- <b>Teorema da aproximação universal de funções</b> .....	56
4.3.2- <b>Limites da aproximação de funções</b> .....	57
4.3.3- <b>Considerações práticas</b> .....	58

4.4- A REDE NEURAL DIRETA MULTICAMADAS PARA A IDENTIFICAÇÃO DAS PRIMEIRAS QUEBRAS.....	60
4.4.1- <b>Arquitetura da rede</b> .....	61
4.4.1.1- Camada de entrada.....	61
4.4.1.2- Camada de saída.....	62
4.4.1.3- Camada oculta.....	62
4.4.1.4- Síntese da arquitetura da rede.....	63
4.5- CONJUNTO DE TREINAMENTO.....	65
4.6- ALGORITMO DE TREINAMENTO.....	65
4.6.1- <b>Termo de momento e adaptação do parâmetro de aprendizagem..</b>	66
<b>5 - RESULTADOS</b> .....	68
5.1- INTRODUÇÃO.....	68
5.2- DADOS SÍSMICOS SINTÉTICOS.....	68
5.2.1- <b>Modelo geofísico</b> .....	69
5.3- TREINAMENTO DA REDE NEURAL DIRETA.....	69
5.4- AVALIAÇÃO DA REDE NEURAL DIRETA.....	71
5.4.1- <b>Avaliação geral da operação da rede</b> .....	71
5.5- AVALIAÇÃO COM SEÇÕES SÍSMICAS COMPLETAS.....	73
5.5.1- <b>Sismograma com um refletor</b> .....	74
5.5.2- <b>Sismograma com dois refletores</b> .....	77
5.5.3- <b>Sismograma com três refletores</b> .....	80
<b>6- CONCLUSÕES</b> .....	84
<b>REFERÊNCIAS BIBLIOGRÁFICAS</b> .....	86

# Lista de Ilustrações

## FIGURAS

Figura 2.1	Terminologia das feições geométricas dos traços sísmicos....	9
Figura 2.2	Registro sísmico.....	11
Figura 2.3	Esquema do levantamento sísmico de refração.....	14
Figura 2.4	Diagrama esquemático do VSP, mostrando ondas ascendentes e descendentes.....	17
Figura 2.5	VSP de fonte fixa.....	18
Figura 2.6	VSP com fonte móvel.....	19
Figura 3.1	Neurônio biológico exemplificado pela célula piramidal, comum no córtex cerebral, mostrando os caminhos do fluxo do sinal (dentritos e axônios), o corpo da célula e, no detalhe, as microcavidades das ligações sinápticas.....	26
Figura 3.2	Modelo de um neurônio artificial típico, onde identifica-se os seus componentes básicos em analogia ao neurônio biológico.....	27
Figura 3.3	Transformação afim realizada pela introdução do termo $\theta_k$ (potencial limite).....	28
Figura 3.4	Função de ativação tipo degrau simples (treshold function)..	29
Figura 3.5	Função de ativação tipo rampa (picewise-linear function)....	30
Figura 3.6	Função de ativação tipo logística, mostrando a variação entre os dois estados (0 e 1), em função do parâmetro $\gamma$ .....	31
Figura 3.7	Rede neural artificial direta, com três camadas, mostrando a completa interconexão entre os neurônios de cada camada consecutiva.....	34
Figura 3.8	Rede neural artificial direta mostrando, através das setas apresentadas sobre as ligações sinápticas, o caminho percorrido pelo sinal, a partir da camada de entrada até a camada de saída.....	38
Figura 3.9	Exemplo de uma arquitetura de rede neural artificial recorrente, mostrando as conexões entre os neurônios e a recorrência representada por $Z^{-1}$ (atraso unitário).....	39
Figura 3.10	Demonstração da generalização realizada pela rede neural direta. A função de mapeamento dos estímulos é suave, produzindo uma boa generalização.....	41

Figura 3.11	Demonstração da generalização realizada pela rede neural direta. A função de mapeamento dos estímulos não é suave, produzindo uma má generalização.....	41
Figura 3.12	Exemplo de uma arquitetura de rede neural artificial direta, onde visualiza-se a entrada do sinal de modo discretizado, a estrutura de camadas e os dois tipos de sinal que percorrem a rede na fase de treinamento.....	43
Figura 3.13	Trajectoria dos sinais no interior da rede durante a fase de treinamento.....	44
Figura 3.14	Rede neural direta com três camadas, onde visualiza-se os pesos sinápticos e o sinal de saída da rede.....	48
Figura 4.1	Exemplo de um registro sísmico contendo apenas dois eventos: a primeira quebra (indicada com seta) e uma reflexão na parte inferior da figura.....	54
Figura 4.2	Traço sísmico, contendo a primeira quebra e uma reflexão, que será usado para compor o conjunto de treinamento da rede.....	60
Figura 4.3	Arquitetura da rede neural para a identificação das primeiras quebras, mostrando também o dado de entrada (seção sísmica) e o dado de saída (seção sísmica contendo somente as primeiras quebras na forma de spikes).....	64
Figura 4.4	Exemplo de um conjunto de treinamento utilizado para a identificação automática das primeiras quebras; onde identifica-se por um círculo, no traço sísmico, a forma de onda representativa das primeiras quebras, a qual formará o estímulo. Na parte inferior da figura, mostra-se a resposta desejada da rede neural.....	67
Figura 5.1	Exemplo do modelo geofísico utilizado para a geração dos sismogramas sintéticos.....	69
Figura 5.2	Exemplos de quatro instantes do treinamento da rede neural direta, mostrando, em azul, o conjunto de estímulos; em verde, o conjunto de respostas desejadas; e, em vermelho, a resposta da rede neural em cada instante do treinamento.....	70
Figura 5.3	Registro de tiro sem ruído. A seta indica o traço sísmico utilizado para o treinamento da rede.....	72
Figura 5.4	Identificação das primeiras quebras através de spikes.....	72
Figura 5.5	Registro de tiro com ruído. A seta indica o traço sísmico que foi utilizado para o conjunto de treinamento.....	73
Figura 5.6	Identificação das primeiras quebras através de spikes.....	73
Figura 5.7	Modelo geofísico utilizado para a geração dos sismogramas sintéticos com um refletor.....	74

Figura 5.8	Sismograma sintético, com um refletor, sem ruído.....	75
Figura 5.9	Primeiras quebras identificadas pela rede neural no sismograma sintético, com um refletor, sem ruído.....	75
Figura 5.10	Sismograma sintético, com um refletor, com ruído.....	76
Figura 5.11	Primeiras quebras identificadas pela rede neural no sismograma sintético, com um refletor, com ruído.....	76
Figura 5.12	Modelo geofísico utilizado para a geração de um sismograma sintético com dois refletores.....	77
Figura 5.13	Sismograma sintético, com dois refletores, sem ruído.....	78
Figura 5.14	Primeiras quebras identificadas pela rede neural no sismograma sintético, com dois refletores, sem ruído.....	78
Figura 5.15	Sismograma sintético, com dois refletores, com ruído.....	79
Figura 5.16	Primeiras quebras identificadas pela rede neural no sismograma sintético, com dois refletores, com ruído.....	79
Figura 5.17	Sismograma sintético, com três refletores, sem ruído.....	81
Figura 5.18	Primeiras quebras identificadas pela rede neural no sismograma sintético, com três refletores, sem ruído.....	81
Figura 5.19	Sismograma sintético, com três refletores, com ruído.....	82
Figura 5.20	Primeiras quebras identificadas pela rede neural no sismograma sintético, com três refletores, com ruído.....	82
Figura 5.21	Exemplos de quatro instantes do treinamento da rede neural direta, mostrando, em azul, o conjunto de estímulos; em verde, o conjunto de respostas desejadas; e, em vermelho, a resposta da rede neural em cada instante do treinamento.....	83

## RESUMO

Apesar do avanço tecnológico ocorrido na prospecção sísmica, com a rotina dos levantamentos 2D e 3D, e o significativo aumento na quantidade de dados, a identificação dos tempos de chegada da onda sísmica direta (primeira quebra), que se propaga diretamente do ponto de tiro até a posição dos arranjos de geofones, permanece ainda dependente da avaliação visual do intérprete sísmico.

O objetivo desta dissertação, insere-se no processamento sísmico com o intuito de buscar um método eficiente, tal que possibilite a simulação computacional do comportamento visual do intérprete sísmico, através da automação dos processos de tomada de decisão envolvidos na identificação das primeiras quebras em um traço sísmico. Visando, em última análise, preservar o conhecimento intuitivo do intérprete para os casos complexos, nos quais o seu conhecimento será, efetivamente, melhor aproveitado.

Recentes descobertas na tecnologia neurocomputacional produziram técnicas que possibilitam a simulação dos aspectos qualitativos envolvidos nos processos visuais de identificação ou interpretação sísmica, com qualidade e aceitabilidade dos resultados. As redes neurais artificiais são uma implementação da tecnologia neurocomputacional e foram, inicialmente, desenvolvidas por neurobiologistas como modelos computacionais do sistema nervoso humano. Elas diferem das técnicas computacionais convencionais pela sua habilidade em adaptar-se ou aprender através de uma repetitiva exposição a exemplos, pela sua tolerância à falta de alguns dos componentes dos dados e pela sua robustez no tratamento com dados contaminados por ruído.

O método aqui apresentado baseia-se na aplicação da técnica das redes neurais artificiais para a identificação das primeiras quebras nos traços sísmicos, a partir do estabelecimento de uma conveniente arquitetura para a rede neural artificial do tipo direta, treinada com o algoritmo da retro-propagação do erro.

A rede neural artificial é entendida aqui como uma simulação computacional do processo intuitivo de tomada de decisão realizado pelo intérprete sísmico para a identificação das primeiras quebras nos traços sísmicos.

A aplicabilidade, eficiência e limitações desta abordagem serão avaliadas em dados sintéticos obtidos a partir da teoria do raio.

## ABSTRACT

In spite of the technologic development happened at seismic prospection, and the significative amount of data with seismic two-dimensional (2D) and three-dimensional (3D) surveys, some process in the seismic interpretation task like the first break picking, remains in a manual version, that still needs an intuitive human intervention.

This dissertation purpose, fill in the seismic processing with the intention to look for an efficient method to enable the computational simulation of the human visual system behavior, through decision process automation involved in first break picking in a seismic trace; looking at to preserv the interpreter intuitive knowledgement to more complex tasks, where your knowledgement will be better profitable.

Neural networks, the most important implementation of neurocomputing systems, were initially developed by neurobiologists as computer models of the neural system in the brain. They differ from conventional computation techniques in their ability to adaptively discriminate or learn through repeated exposure to examples, their tolerance to data component failure and their robustness in the presence of high noise levels. This computing technology provide some techniques that can reduce the labor intensive aspects of the first break picking, maintaining the quality and reliability of the results.

The method here presented is an application of an artificial neural network computational process, known as feedforward multilayer perceptron trained with the error back-propagation algorithm; from the establishment of a convenient neural network architecture and learning set that make possible its application over seismic data. This method is a computational simulation of seismic interpreter decision intuitive process for first break picking in seismic traces.

The applicability, efficiency and limitations of this approach will be appraised in synthetic data obtained starting out the ray theoretical method.

# 1- INTRODUÇÃO

Na grande maioria dos levantamentos sísmicos terrestres encontramos uma enorme variação litológica dos depósitos superficiais de sedimentos inconsolidados (argila, areia, cascalho, etc...) e grandes alterações nas elevações topográficas da superfície que afetam, prejudicialmente, os tempos de trânsito das ondas sísmicas. Na prospecção sísmica, esses estratos superficiais recebem a designação geral de zona de intemperismo ou zona de baixa velocidade (ZBV).

Nesta dissertação tratamos, com particular interesse, a geometria do caminho do raio de uma frente de onda que se propaga do ponto de explosão diretamente para a posição dos geofones, através da zona de baixa velocidade, transformando-se no primeiro sinal representativo da chegada de uma frente de onda no traço sísmico resultante.

À este particular evento, no traço sísmico, dá-se o nome de primeira quebra (first break), o qual representa o tempo necessário para que a onda direta percorra a distância de separação entre a fonte e o arranjo de geofones com a velocidade característica da zona de intemperismo.

A tecnologia do processamento sísmico tem sido capaz de manter, mesmo com o aumento do volume de dados, a sua eficiência através de melhorias na capacidade e velocidade de processamento dos computadores (McCORMACK et al., 1993). Entretanto, há três áreas no processamento de dados sísmicos que não têm sua tecnologia computacional totalmente explorada e permanecem com suas análises na forma visual:

- ◆ Determinação das primeiras quebras.
- ◆ Edição de traços sísmicos com ruído.
- ◆ Análise de velocidades.

Apesar de todo o esforço técnico-científico para melhorar o tempo de processamento dos dados sísmicos, a identificação das primeiras quebras (first break picking) continua a ser realizada de forma manual, valendo-se do esforço visual de um intérprete sísmico.

O objetivo desta dissertação, insere-se no processamento sísmico com o intuito de buscar um método eficiente, tal que possibilite a simulação computacional do comportamento visual do intérprete sísmico, através da automação dos processos de tomada de decisão envolvidos na identificação das primeiras quebras em um traço sísmico; visando, em última análise, preservar o

conhecimento intuitivo do intérprete para os casos complexos, nos quais o seu conhecimento será, efetivamente, melhor aproveitado.

A importância da identificação automática das primeiras quebras, nos traços sísmicos monocanal ou multicanal, pode ser visualizada através do processamento para a realização das correções estáticas. O método mais conhecido para este fim é o de Hagedoorn (HAGEDOORN, 1959), ainda hoje empregado. O progresso computacional inserido neste método, desde então, foi, somente, o de possibilitar que o intérprete realize a identificação visual das primeiras quebras diretamente na tela do computador.

Recentes descobertas na tecnologia neurocomputacional produziram técnicas que buscam simular os aspectos qualitativos envolvidos nos processos visuais interpretativos da identificação das primeiras quebras, enquanto é mantida a qualidade e aceitabilidade dos resultados (McCORMACK et al., 1993). As redes neurais artificiais são uma implementação dos sistemas neurocomputacionais e foram, inicialmente, desenvolvidas por neurobiologistas como modelos computacionais do sistema nervoso humano. Elas diferem das técnicas computacionais convencionais pela sua habilidade em adaptar-se ou aprender através de uma repetitiva exposição a exemplos, pela sua tolerância à falta de um componente dos dados e pela sua robustez na presença de ruídos (HAYKIN, 1994).

O método aqui apresentado, baseia-se na aplicação da técnica das redes neurais artificiais; a partir do estabelecimento de uma conveniente arquitetura e de um conjunto de treinamento para uma rede neural artificial do tipo direta, treinada com o algoritmo da retro-propagação do erro (RUMELHART et al., 1986; McCORMACK et al., 1993; ANDRADE & LUTHI, 1997), entendida aqui como uma simulação computacional do processo intuitivo de tomada de decisão do intérprete sísmico para a identificação das primeiras quebras nos traços sísmicos.

A aplicabilidade, eficiência e limitações desta abordagem serão avaliadas em dados sintéticos obtidos a partir da teoria do raio (CERVENY, 1987).

Esta dissertação está estruturada na seguinte forma:

No Capítulo 2 apresenta-se uma visão geral dos métodos sísmicos e das técnicas de aquisição do dado sísmico.

No Capítulo 3 apresenta-se um resumo das redes neurais artificiais, mostrando os principais tipos de redes, com ênfase na rede neural direta multicamadas - escolhida como o

modelo de rede a ser utilizado nesta dissertação. Avalia-se todo o processamento sofrido pelo sinal no interior da rede e mostra-se a dedução matemática do algoritmo de retro-propagação do erro (error-backpropagation).

No Capítulo 4, mostra-se a metodologia desenvolvida nesta dissertação para a automação do processo de tomada de decisão do intérprete sísmico para a identificação das primeiras quebras nos traços sísmicos, por meio de uma rede neural do tipo direta multicamadas, treinada com o algoritmo da retro-propagação do erro. Discute-se amplamente a concepção da arquitetura da rede apresentada e do seu conjunto de treinamento.

No Capítulo 5, avalia-se o desempenho da arquitetura de rede apresentada na identificação das primeiras quebras, em sismogramas sintéticos, na forma de registros de tiro, com a fonte situada em duas posições do modelo. Avalia-se ainda o comportamento da rede com dados ruidosos para ambos os tipos de seção.

Nas Conclusões, destaca-se a adequabilidade da arquitetura de rede apresentada, bem como uma introdução do método de avaliação do processamento com as redes neurais, apresentando possíveis futuras aplicações da metodologia utilizada.

## 2- MÉTODOS SÍSMICOS

### 2.1- INTRODUÇÃO

A prospecção geofísica, com o método sísmico, descende diretamente da sismologia, que trata do estudo das grandes estruturas geológicas do interior da Terra em escala planetária. A sismologia trata do registro, em estações sismológicas espalhadas por todos os continentes, dos tempos de trânsito das ondas sísmicas, geradas naturalmente pela energia liberada durante os terremotos. A interpretação dessas informações subsidia os geólogos para um melhor conhecimento da estrutura do interior do planeta.

A prospecção sísmica trata da utilização das ondas sísmicas, artificialmente geradas, com o intuito de obter um maior conhecimento geológico da subsuperfície em escala local. A técnica básica da exploração sísmica inicia-se na geração artificial de ondas sísmicas, a partir de fontes de energia controlada e móveis (dinamite, vibroseis, air gun). A energia dessas fontes dá origem às ondas sísmicas que se propagam pela subsuperfície, sofrendo todos os fenômenos da mecânica ondulatória e, após um período de tempo, retornam à superfície onde são registradas por um conjunto de receptores (geofones) regularmente dispostos. Os movimentos da superfície, que são registrados pelos geofones, são transformados em sinais elétricos, que podem ser armazenados na forma digital ou analógica, para o seu posterior processamento e interpretação.

A partir dos conhecimentos dos tempos de trânsito, para os vários geofones, e a velocidade das ondas sísmicas, pode-se reconstruir os caminhos percorridos pelas mesmas. A informação geológico-estrutural é obtida, principalmente, através da trajetória das ondas sísmicas.

O objetivo da exploração sísmica é deduzir informações sobre as rochas, especialmente a respeito das atitudes das camadas, a partir dos tempos de chegadas observados e, num sentido restrito, a partir de variações da amplitude, frequência e forma de onda.

### 2.2 – ONDAS ACÚSTICAS

A propagação de ondas mecânicas pode ser bem estudada a partir da equação da onda acústica em um meio com densidade constante e velocidade variável. No caso da prospecção sísmica, a subsuperfície é representada por um modelo tridimensional, com o vetor posição expresso por  $\vec{x}=(x,y,z)$  e a fonte sísmica, considerada pontual, expressa por  $\delta(\vec{x}-\vec{\xi})\delta(t-\tau_0)$ .

O vetor deslocamento da partícula  $\vec{u}(\vec{x}, t)$ , em um meio acústico, onde a velocidade de propagação da onda é representada por  $c=c(\vec{x})$ , satisfaz a equação denominada de equação da onda acústica, dada por

$$\nabla^2 \vec{u}(\vec{x}, t) - \frac{1}{c^2(\vec{x})} \frac{\partial^2 \vec{u}(\vec{x}, t)}{\partial t^2} = -\delta(\vec{x} - \vec{x}_s) \delta(t - t_0) \quad (2.1)$$

Na equação acima assume-se que  $x_s$  e  $t_0$  são conhecidos. A função  $\vec{u}(\vec{x}, t, \vec{x}_s, t_0)$  é a função de Green. Os seus dois primeiros argumentos caracterizam os receptores, na forma

$\vec{x}$  - vetor posição do geofone.

$t$  - variável independente temporal.

Os dois últimos argumentos caracterizam a fonte, onde

$\vec{x}_s$  - vetor posição da fonte.

$t_0$  - tempo instantâneo na fonte.

O termo  $\delta(\vec{x} - \vec{x}_s) \delta(t - t_0)$  representa a fonte pontual no espaço e no tempo.

Considerando que a fonte sísmica é representada pelo impulso localizado em  $\vec{x} = \vec{x}_s$ , no instante  $t = t_0$ , a solução desta equação, no domínio da frequência, é dada por (PSENCIK, 1994)

$$u(\vec{x}, t; \vec{x}_s, t_0) = \frac{1}{4\pi r} \delta\left(t - t_0 - \frac{r}{c}\right) \quad (2.2)$$

onde  $r = |\vec{x} - \vec{x}_s|$ , corresponde à distância entre fonte e receptor.

A função de Green representa a resposta do meio a uma fonte impulsiva. A sua utilização possibilita, ainda, a obtenção de soluções da equação acústica com termos representativos da fonte sísmica numa forma mais complexa (PSENCIK, 1994).

## 2.3 - O MÉTODO SÍSMICO DE REFLEXÃO

Este método busca o mapeamento da subsuperfície através da utilização dos tempos requeridos por uma onda mecânica, gerada a partir de uma fonte sísmica artificial localizada próximo à superfície, retornar até os instrumentos detectores, após ser refletida em profundidade.

O movimento do terreno nas proximidades do geofone é transformado em um sinal elétrico que pode ser registrado na forma digital ou analógica. Os eventos sísmicos de reflexão, que são as chegadas aos geofones da energia sísmica refletida nas interfaces em subsuperfície, são identificados nos registros. Os tempos de chegada (o intervalo entre o instante do tiro e a chegada da energia no grupo de geofones, também conhecido como tempo de trânsito) desses eventos são medidos em vários grupos de geofones. A localização e atitude da interface que deu origem à cada evento de reflexão são, então, calculadas a partir dos tempos de trânsito e da velocidade. Deste modo, o dado registrado (seção sísmica) constitui-se de um conjunto de traços, onde cada traço mostra o movimento sofrido por cada geofone ao longo do tempo.

Os dados registrados pela sísmica de reflexão são, usualmente, processados para a atenuação dos ruídos, correção estática, migração e etc, de tal modo a poderem ser grafados na forma apropriada para a sua interpretação geológica.

As seções sísmicas são produzidas a partir de um conjunto de traços sísmicos. Um traço é constituído por um conjunto de dados que são recebidos por um grupo de geofones, a partir da detonação de uma fonte sísmica. Ele é visualizado na forma de uma linha ondulada. Na sua forma convencional, os dados são representados como ondulações dessa linha.

A expressão traço sísmico teve sua origem na sismologia, onde o registro do movimento da terra é realizado pelo traçamento de uma linha. Tomando-se um traço sismológico horizontal, com seu início na esquerda e os tempos posteriores à direita, as ondulações serão para cima e para baixo. As ondulações para cima são chamadas de *pico* e as para baixo de *calha* (COFFEEN, 1978), como mostrado na Figura 2.1. A convenção é mantida, mesmo quando o traço é posto na vertical. Então, os *picos* são para a direita e as *calhas* para a esquerda, continuando com a mesma denominação. Um traço sem ondulações é dito um traço morto. Como um traço sísmico mostra o registro do movimento da terra, o *pico* representa um movimento para cima e a *calha* um movimento para baixo. A direção dos traços não é completamente unificada; logo, traços de diferentes operações de campo podem diferir no seu modo de apresentação, sendo dito que eles possuem diferentes polaridades.

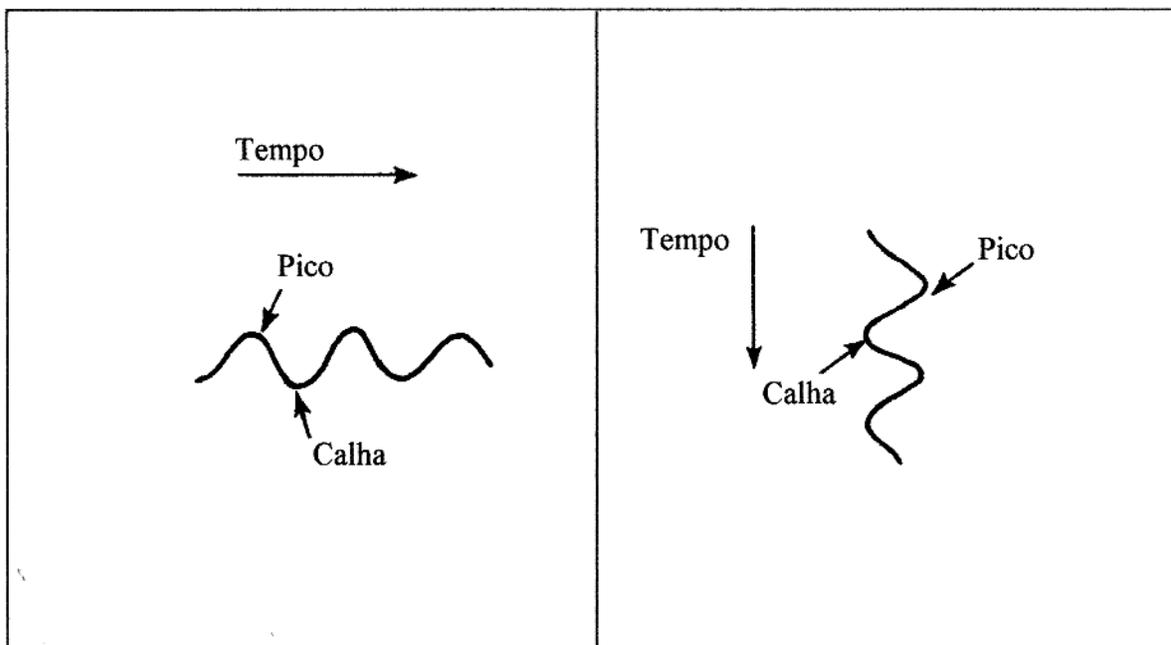


Figura 2.1: Terminologia das feições geométricas dos traços sísmicos.

Um registro de tiro é uma composição de todos os traços sísmicos gerados a partir de um mesmo tiro (detonação da fonte sísmica). Esta forma de apresentação do traço sísmico foi utilizada por muitos anos como única fonte de interpretação sísmica, antes do desenvolvimento da construção das seções sísmicas. Atualmente, suas principais utilizações são o monitoramento da qualidade dos dados de campo e definição dos tempos das primeiras quebras.

Os registros de tiro são normalmente apresentados com 24, 48 ou mais traços sísmicos postos lado a lado, sendo que cada traço é representativo do agrupamento de geofones que estão conectados a um canal de registro. Os traços estão dispostos de forma que os *picos* e as *calhas* de um traço ficam sobrepostos com os dos outros traços.

Ainda no registro de tiro são apresentadas as linhas de tempo, as quais normalmente são marcadas na horizontal a cada 4 ms, cuja finalidade é fornecer uma indicação do tempo decorrido no interior da janela temporal de registro, a qual é fixada para cada levantamento de campo em particular.

Analisando de cima para baixo o registro sísmico (Figura 2.2), a partir dos primeiros tempos para os posteriores, os traços serão, inicialmente, quase retos, somente com pequenas ondulações causadas pela ação do vento ou outras leves perturbações sobre os arranjos de geofones.

A primeira ondulação relevante, verificada no registro, é chamada tempo de quebra (time break). Ela aparece no traço sísmico na forma de uma onda abrupta, referindo-se ao instante de tempo no qual a fonte sísmica é detonada. O tempo de quebra é usado como ponto de partida para a medição dos tempos de todos os outros eventos sísmicos contidos no traço. O tempo de quebra não corresponde, evidentemente, a um evento sísmico ou se refere a um particular grupo de geofones; ele é arbitrariamente posto no traço como forma de marcar visualmente os tempos dos eventos relevantes.

A próxima ondulação, a primeira representativa de um evento sísmico, recebe o nome de primeira quebra (first break). Normalmente, posicionam-se no registro sísmico ao longo de linhas ou segmentos de linhas retas, indicando a velocidade da camada próxima à superfície ao longo da qual as ondas sísmicas propagam-se diretamente da fonte aos grupos de geofones.

Após as primeiras quebras aparecem as reflexões, na forma de ondulações, que se alinham de traço para traço ao longo do registro. A energia refletida faz um percurso diferente do realizado pelas primeiras quebras, o percurso feito pela reflexão é curvo, sendo para baixo com o aumento da distância da fonte ao geofone. As reflexões mais rasas curvam-se consideravelmente mais que as reflexões mais profundas.

Um grande número de traços é combinado para produzir o registro da seção. Normalmente, a seção é obtida a partir da justaposição dos registros. Uma seção de registro moderna, seção sísmica ou apenas seção, é mais sutilmente construída, obtendo seus dados a partir de um número de registros ajustados e combinados para produzir, como resultado, algo notavelmente parecido com um corte transversal da subsuperfície terrestre. Uma seção possui centenas de traços lado a lado, sobrepondo-se. O número de pontos de tiro são impressos ao longo do topo de cada seção. Eles podem ser ajustados em número no mapa, para local os pontos de tiro da seção no solo.

As reflexões são vistas através da seção, com uma curvatura diferente da apresentada no registro sísmico, a qual é removida durante o processamento sísmico.

As reflexões em uma seção mostram mais ou menos as camadas de rocha no subsolo como numa seção geológica. Com limitações, as camadas de rocha vão para cima, para baixo, estão quebradas e etc, conforme indicado nas reflexões.

Na fase de interpretação, sobre a seção sísmica, são marcadas todas as reflexões. Após, realiza-se a interpolação dessas reflexões entre as diversas seções, mantendo-se o cuidado particular dos pontos de amarração (interseção de linhas sísmicas). Os tempos dessas reflexões são impressos no plano e contornados, mostrando a configuração espacial das camadas rochosas, a partir da qual a locação para a perfuração de poços será deduzida.

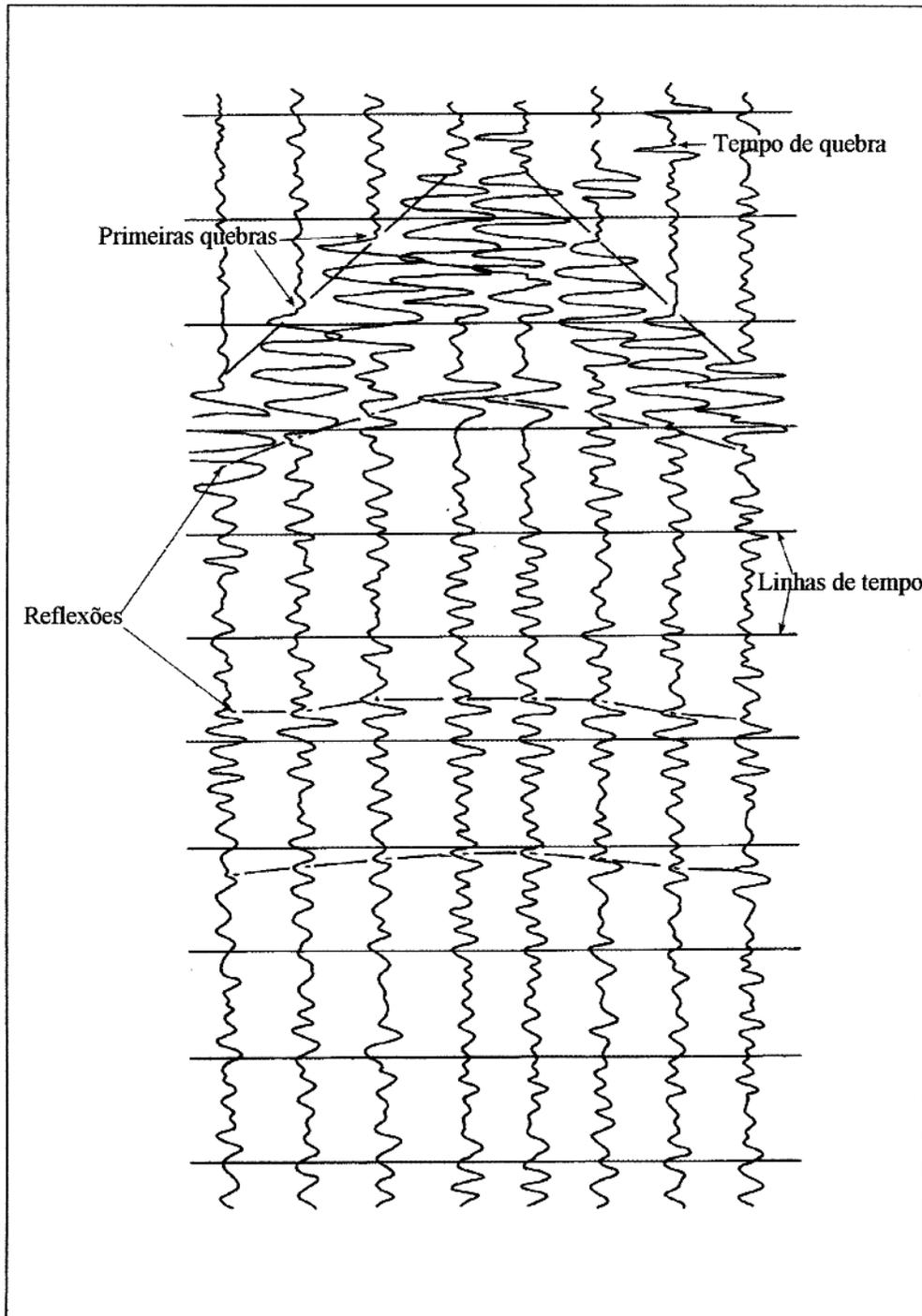


Figura 2.2: Registro sísmico.

## 2.4 - O MÉTODO SÍSMICO DE REFRAÇÃO

Os fundamentos do método de refração sísmica foram elaborados para a sismologia dos terremotos muito antes de se pensar na prospecção sísmica. Aplicando estes conhecimentos, os sismólogos puderam determinar a estrutura interna da Terra fazendo uso dos tempos requeridos pelas ondas sísmicas para deslocar-se desde sua origem até estações registradoras situadas em pontos distantes. Na prospecção por refração, a estratificação do subsolo detém-se a uma escala muito menor, utilizando os tempos de trânsito das ondas ocasionadas por explosões provocadas na superfície. Em princípio, só se aproveitarão para estes fins os tempos de primeiras chegadas, mas, ultimamente, também utiliza-se os acontecimentos posteriores.

O método de refração sísmica é especialmente valioso para reconhecimentos em áreas cuja estrutura tem grande relevo e de onde tem, pelo menos, uma camada de referência de alta velocidade recoberta por formações de baixa velocidade.

No campo, as prospecções por refração são mais completas que as de reflexão porque o trabalho está disseminado em uma distância muito maior. Na prospecção por reflexão, a explosão e os detectores nunca estão mais separados que a profundidade da formação estudada e, em certas ocasiões, muito mais próximos. Na prospecção por refração, a separação é muito maior que a profundidade em que se encontra a interface.

A refração sísmica é governada pela Lei de Snell

$$\frac{V1}{V2} = \frac{\text{sen } \alpha}{\text{sen } \beta} \quad (2.3)$$

onde  $V1$  e  $V2$  são as velocidades das ondas sísmicas acima e abaixo de uma interface de separação entre duas camadas de litologias diferentes. A refração dá-se quando  $\alpha$  é igual a  $90^\circ$ , neste caso  $\beta$  é chamado de ângulo crítico. Nessa situação, a onda refratada propaga-se de modo tangente à interface.

A mais simples forma do método de refração é a do tiro reverso, na qual uma seção interpretada poderia ser bem correlacionada com uma seção sísmica de reflexão, com a vantagem de que no método de refração os horizontes são definidos em termos de profundidade e não no tempo duplo de reflexão.

Os métodos de análise e processamento dos dados de refração podem ser bastante complicados, a menos que as condições geológicas sejam bastante simples, por exemplo: camadas horizontais, homogêneas e isotrópicas e ausência de inversão de velocidade.

Os locais para a aplicação do método devem ser escolhidos, criteriosamente, de modo a não se afastarem muito das condições ideais, principalmente em trabalhos de reconhecimento ou em situações onde o método é utilizado para obter informações sobre a velocidade e controle da profundidade no auxílio ao método de reflexão.

Somente o caso mais simples será tratado aqui, no qual duas camadas planas com velocidades sísmicas  $V_1$  e  $V_2$  estão superpostas ao embasamento com velocidade  $V_3$ . Assumimos a posição do tiro direto em  $O$ , com o grupo de receptores colocados em  $P$ . Os receptores receberão as ondas sísmicas refratadas com velocidade  $V_2$  ou com velocidade  $V_3$ , dependendo de sua distância em relação ao ponto de tiro. Um típico gráfico tempo-distância é mostrado na Figura 2.3. A partir dele, é possível obtermos informações sobre a velocidade e a espessura das camadas. A inclinação das retas é igual ao inverso da velocidade.

Um perfil de refração como descrito acima, por si só não detecta a presença de eventos mergulhantes ou quaisquer outros desvios das condições de camada plano-horizontal. Por isso procede-se ao tiro reverso, trocando-se as posições de fonte e receptor. Nas condições ideais, deve-se ter um gráfico tempo-distância exatamente simétrico ao do tiro direto, como mostrado na Figura 2.3, em linha pontilhada.

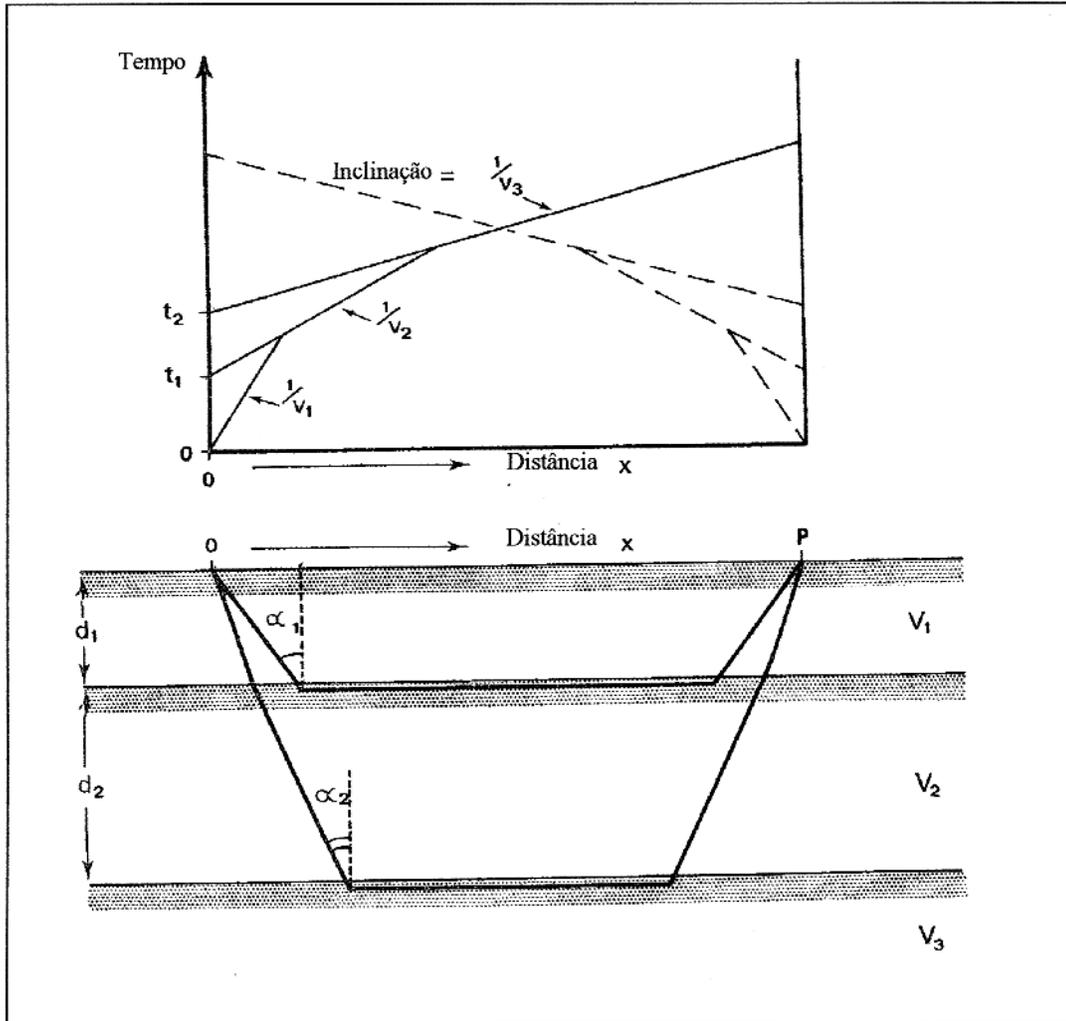


Figura 2.3: Esquema do levantamento sísmico de refração.

## 2.5 – SÍSMICA DE POÇO

A sísmica de poço teve a sua origem na prospecção sísmica de superfície, a partir da utilização de uma fonte sísmica convencional posicionada na superfície e grupos de geofones dispostos semi-verticalmente no interior de um poço, para o registro dos tempos de trânsito. Com o incremento da perfuração de poços direcionais e poços horizontais, a utilização da sísmica de poço tem crescido consideravelmente no âmbito das companhias petrolíferas.

De uma forma geral os equipamentos da sísmica de poço são os mesmos da sísmica de superfície. A fonte de ondas sísmicas artificiais normalmente é um *air gun*, o qual é colocado no interior de um tanque com água com poucos metros de profundidade, nas proximidades da boca do poço. O equipamento de registro das ondas sísmicas é formado por um arranjo com um ou mais hidrofones, os quais são montados em um cabo flexível, que tem como objetivo a redução do ruído e tornar-se um meio comum para o deslocamento de todo o arranjo ao longo das paredes do poço. Os hidrofones são equipamentos sensíveis às ondas mecânicas no intervalo das frequências sísmicas, a partir de poucos Hertz até 1000 Hz. Os arranjos mais antigos eram constituídos, simplesmente, por hidrofones unidirecionais; enquanto os mais atuais, caracterizam-se pela utilização de hidrofones tridirecionais alinhados ao longo dos três eixos principais do poço.

Para a pesquisa de levantamento de poço, a medida de interesse é o tempo de trânsito a partir da superfície até a localização do hidrofone, ou seja, as primeiras quebras. Os tempos de trânsito registrados em várias profundidades são usados para calibrar os traços sísmicos em termos da sua conversão tempo/profundidade. Em poços direcionais, torna-se necessário uma pesquisa para corrigir a profundidade verdadeira abaixo da superfície. Em poços que atravessam camadas inclinadas, os resultados do *dipmeter* podem dar-nos a informação necessária para misturarmos a sísmica de superfície com a melhoria da migração das linhas sísmicas.

Para o perfil sísmico vertical – VSP (vertical seismic profile), mostrado na Figura 2.4, o princípio básico é similar, excetuando que o registro do tempo é muito maior, tal que, além das primeiras quebras, todas as ondas refletidas, incluindo as múltiplas, são registradas. A ferramenta é movida em incrementos de profundidade poço acima relativamente constantes. Normalmente, os incrementos de profundidade são da ordem de 10 à 50 metros, com 1 à 5 tiros em cada profundidade. Desta maneira, as ondas descendentes (downgoing waves) podem ser separadas

das ondas ascendentes (upgoing waves). Entre as ondas descendentes encontramos as primeiras quebras e as múltiplas, cujo último refletor está acima da posição do geofone. Como a posição do geofone é movida poço acima, elas chegam primeiro até que o geofone cruze o horizonte de reflexão e elas tornem-se ondas ascendentes. Entre as ondas ascendentes estão todas as reflexões, incluindo as múltiplas, cujo refletor está abaixo da posição do geofone. Para o perfil sísmico vertical – VSP, interessa somente as ondas ascendentes, pois elas correspondem às reflexões registradas pela sísmica de superfície. Os trens de onda (wavetrains) registrados são, então, separados em ondas descendentes e ascendentes. As técnicas de filtragem na frequência e de deconvolução são aplicadas para melhorar a resolução espacial e eliminar ruído. Normalmente, somente frequências entre 20 e 100 Hz são mantidas. O trem de onda descendente é usado para calcular o operador de deconvolução para o trem de onda ascendente. Em um poço vertical, o VSP nos dá informações similares à linha sísmica de superfície para a localização do poço, mas com uma resolução vertical consideravelmente melhor porque as camadas rasas são cruzadas somente uma vez e, então, temos menos filtragem pela terra que na sísmica de superfície. O VSP também contém um registro preciso dos intervalos de tempo de trânsito da conversão tempo/profundidade. Uma característica particularmente atrativa do VSP é que ele contém informações abaixo da profundidade total do poço. Isto pode ser usado para a predição de uma estrutura reservatório abaixo do fim do poço.

Um tipo mais recente de VSP é o offset seismic profile (OSP). Na fonte fixa VSP, a fonte é movida para uma certa distância afastada do poço. Os pontos de reflexão de um horizonte de interesse mudam agora conforme a posição do geofone no poço é mudada. A proximidade do geofone está para o horizonte assim como a proximidade do ponto de reflexão estará para o poço. Então, o refletor poderá ser mapeado fora do poço. No OSP de caminho (walkaway OSP), a fonte é movida continuamente para fora, enquanto o geofone continua numa posição estacionária. Os pontos de reflexão movem-se agora para fora do poço, assim como a fonte é movida para fora do mesmo, dando novamente uma zona imageada do refletor de interesse. Uma variação adicional desta técnica é o perfil sísmico multi-offset, onde a fonte é movida radialmente para fora do poço ao longo de várias linhas, normalmente em padrão estrela. Desta maneira, uma imagem tridimensional da estrutura da subsuperfície nas vizinhanças do poço é

obtida. Em poços direcionados, o OSP pode ser obtido com o movimento simultâneo de fonte e geofones, tal que eles recaiam exatamente em cima de cada um em toda posição de tiro.

As técnicas do VSP (Figura 2.5) e OSP (Figura 2.6) são uma nova tecnologia emergente, as quais preenchem as lacunas entre a sísmica de superfície e a perfilagem de poço. Elas representam os sismogramas reais em oposição aos sismogramas sintéticos, também chamados geogramas, os quais são calculados a partir do perfil de poço sônico e medições de densidade, que são deixadas para segundo plano, com melhorada aquisição e tecnologia de processamento disponível para os VSPs e OSPs. Suas principais desvantagens são custos altos e um grande volume de dados.

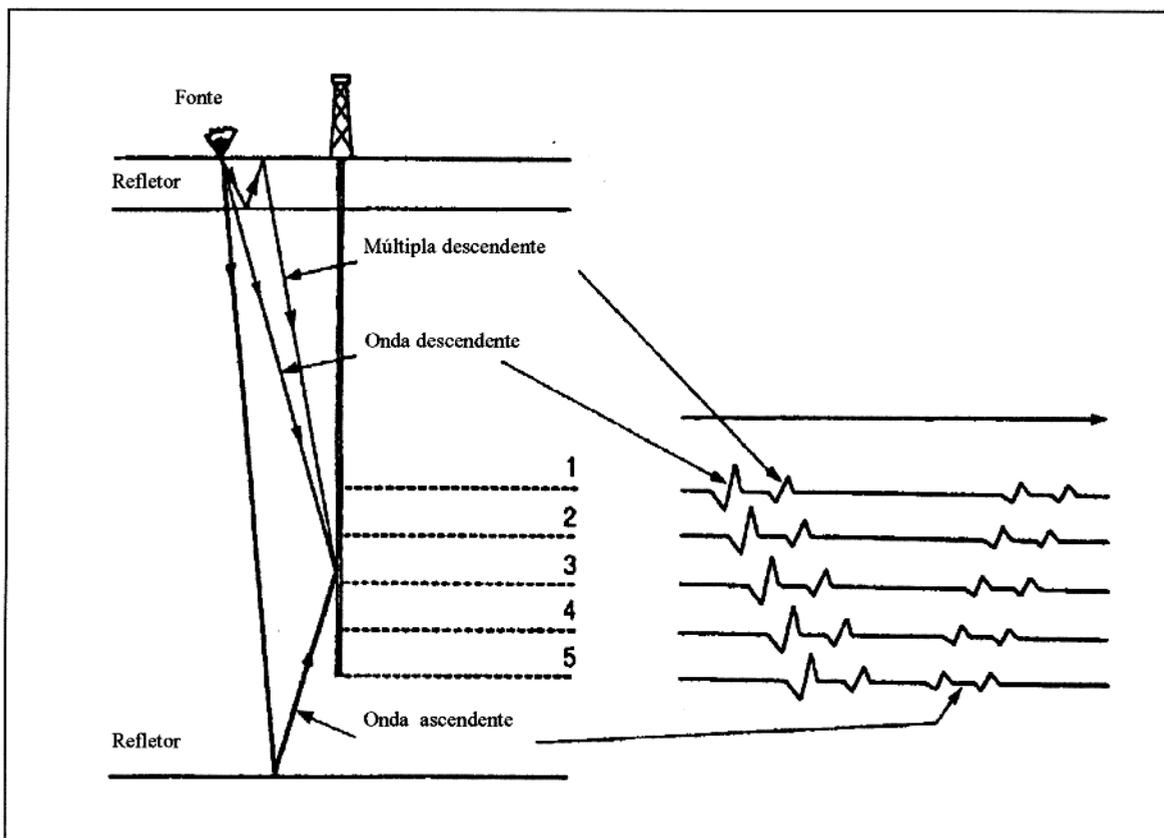


Figura 2.4: Diagrama esquemático do VSP, mostrando ondas ascendentes e descendentes.

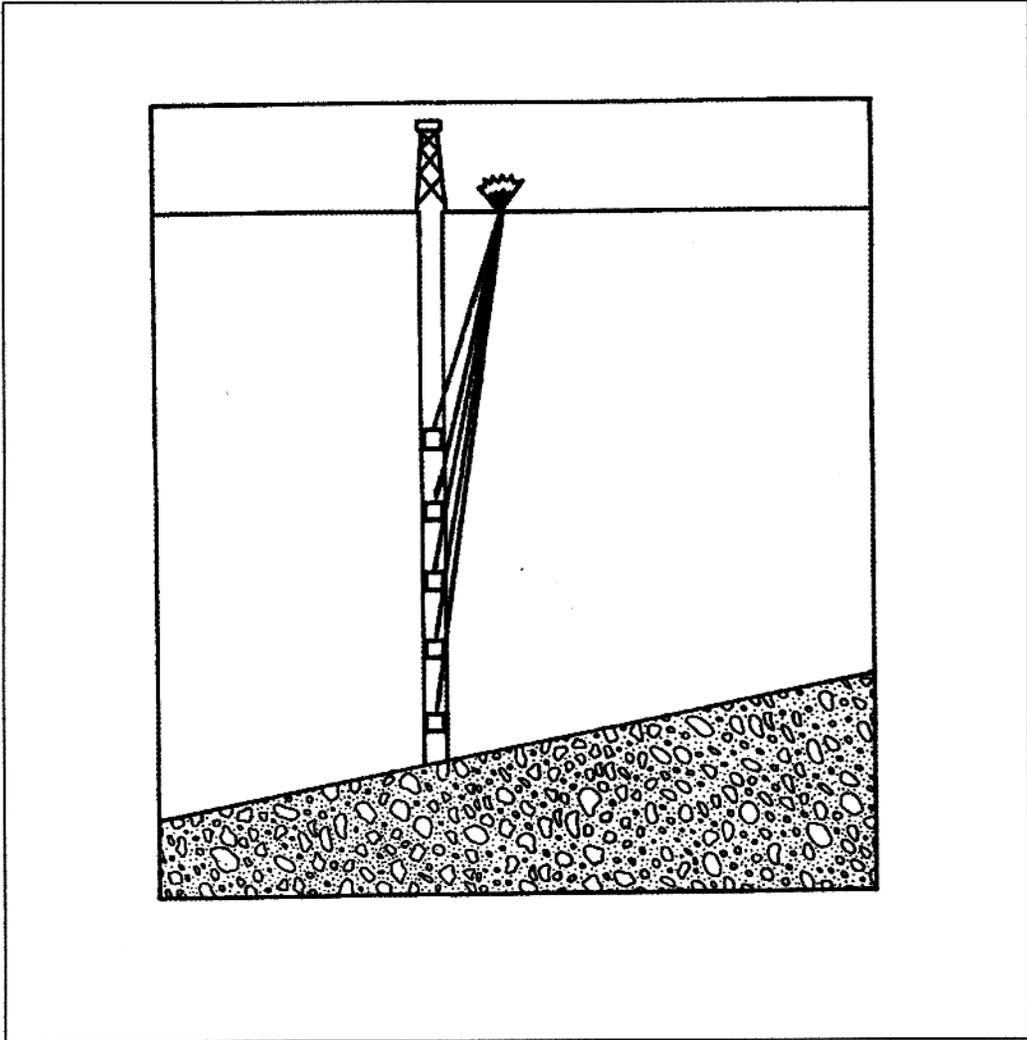


Figura 2.5: VSP de fonte fixa.

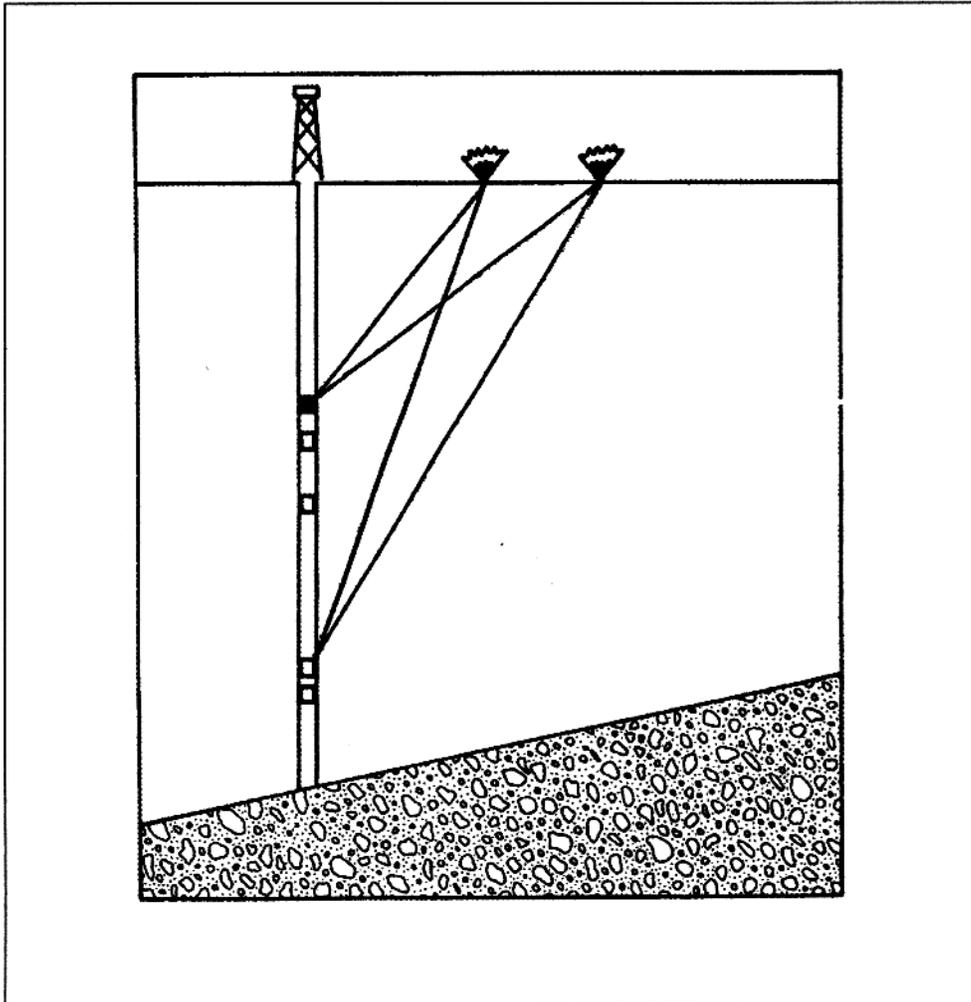


Figura 2.6: VSP com fonte móvel.

## 3- REDES NEURAIS ARTIFICIAIS

### 3.1-INTRODUÇÃO

Por várias décadas a ciência tem alcançado um considerável progresso no desenvolvimento de máquinas ditas “inteligentes”. Essas máquinas seriam destinadas a executar tarefas com a habilidade de poder aperfeiçoar, no tempo, as suas próprias operações. Entre as tecnologias empregadas para alcançar esse objetivo pode-se citar a matemática biofísica, a cibernética, a automática, a biocibernética, a teoria geral de sistemas, os sistemas auto-organizativos, a inteligência artificial, a ciência cognitiva e os sistemas neurais artificiais, também conhecidos por redes neurais artificiais.

A criação de algoritmos para os sistemas neurais é motivada pelo reconhecimento de que o cérebro humano trabalha de forma inteiramente diferente que os computadores digitais convencionais, operando sob algoritmos sequenciais (HAYKIN,1994). Assim, a idéia fundamental das redes neurais artificiais é a de simular, por meio de algoritmos, o comportamento do cérebro humano, quando este trabalha na solução de problemas triviais, mas que apresentam um tratamento bastante complexo e oneroso nos computadores convencionais.

A história das redes neurais artificiais tem com marco inicial, o ano de 1943, quando Warren Mcculloch, um neurofisiologista, e Walter Pitts , um matemático de apenas 18 anos de idade, escreveram um artigo intitulado “*about how neurons might work*” (McCulloch & PITTS, 1943). Eles modelaram uma rede neural simples com circuitos elétricos. Dessa forma, eles conseguiram montar o primeiro modelamento matemático de uma rede neural simples que era capaz de processar dados, mas não era capaz de aprender com estes dados.

Seis anos depois, em 1949, uma outra grande contribuição foi dada para o crescimento das redes neurais artificiais por Donald Hebb , quando tornou público o livro denominado “*The organization of Behavior*” (HEBB, 1949), onde o mesmo propôs um esquema de aprendizagem para atualização das conexões entre neurônios, resultando na elaboração da regra que hoje é conhecida como “regra de aprendizagem de Hebb”.

Na década de cinquenta, com os aperfeiçoamentos em hardware e software, os computadores de segunda geração foram construídos (MINSKY, 1954). Deu-se também nessa época a primeira construção física de uma rede neural artificial, por Edmonds & Minsky. Ainda

nessa década, McClelland defende a idéia de que a representação do conhecimento no cérebro faz-se de maneira distribuída. Deste modo, a responsabilidade de armazenar um determinado conceito não fica colocada a cargo de uma única unidade, mas sim de um grupo de unidades (neurônios).

Pesquisadores do laboratório de pesquisa da IBM conduziram uma simulação em software de um modelo de rede neural baseado no trabalho de Hebb, em 1950. A primeira tentativa falhou, mas com a colaboração de Hebb e outros, adaptações foram feitas com sucesso.

Nos anos seguintes, John Von Neumann escrevendo material para seu livro *The Computer and the Brain* (VON NEWMANN, 1958), faz sugestões de imitar-se as funções de um simples neurônio, usando relés de telégrafo e válvulas eletrônicas. Frank Rosenblatt (ROSENBLATT, 1958) começou um trabalho sobre o *Perceptron*, um modelo de rede neural que prometeu bastante, na medida em que era capaz de aprender padrões e generalizar a partir dos padrões aprendidos, pela modificação dos pesos das conexões. Esta idéia foi uma das maiores contribuições à teoria de redes neurais artificiais, servindo como base para os algoritmos de aprendizagem que hoje são estudados.

Em 1959, Bernard Widrow & Marcian Hoff (Stanford) desenvolveram os modelos de rede ADALINE e MADALINE (Multiple ADaptive LINEar Elements) (WIDROW & HOFF, 1960). Esta foi a primeira rede neural no mundo aplicada a um problema real – filtros adaptativos para eliminar ecos (ruídos) sobre os sinais telefônicos. Eles introduziram uma nova regra de aprendizagem chamada *Widrow-Hoff learning rule* (WIDROW & HOFF, 1966). Essa regra minimizava o somatório dos quadrados dos erros durante o treinamento envolvendo classificação de padrões. Mais tarde a adaline foi usada com sucesso para o processamento adaptativo de sinais (WIDROW et al., 1963; WIDROW, 1973; WIDROW et al., 1975; WIDROW & STEARNS, 1985), sistemas de controle (WIDROW & SMITH, 1964; WIDROW, 1988), e sistemas adaptativos de antena (WIDROW, MANTEY & GRIFFITHS, 1967).

Steinbuch, em 1961, foi um dos primeiros pesquisadores a desenvolver o método da codificação da informação em um sistema neural artificial “crossbar” (tipo de sistema de comutação de linhas que utiliza computadores de barras cruzadas). O método foi aplicado a problemas, tais como: reconhecimento de caracteres altamente distorcidos escritos à mão,

diagnósticos de falhas mecânicas e etc. (STEINBUCH & PISKE, 1963; STEINBUCH & ZENDEH, 1963).

Apesar dos vários sucessos obtidos e entusiasmo dos pesquisadores no início e meados da década de 60, houve um esfriamento à corrida das redes neurais artificiais. Com a publicação do livro *Perceptrons* por Minsky & Papert (MINSKY & PAPERT, 1969), aumentaram-se as dúvidas a respeito do potencial das redes neurais. Nesta obra, os autores provam, formalmente, a incapacidade dos perceptrons de aprenderem funções como, por exemplo, um simples XOR. Mostram também a existência de máximos e mínimos locais que poderiam ser encontrados durante o processo de aprendizagem e que certamente conduziriam à um processo de aprendizagem insatisfatório.

Mesmo com o ritmo desacelerado das pesquisas em redes neurais artificiais, a área de inteligência artificial continuou sendo alvo dos pesquisadores. Um dos maiores pesquisadores e que mais influenciou nas pesquisas de redes neurais artificiais, da década de 60 até os dias atuais, é Stephen Grossberg (Boston University). Em 1964, Stephen Grossberg estudou exaustivamente os processos psicológicos (mente), biológicos (cérebro) e fenômenos do processamento da informação humana, unindo-os dentro de uma teoria unificada (GROSSBERG, 1964). Ele divulgou vários estudos sobre o paradigma das redes neurais artificiais, analisando formalmente aspectos do modelo de aprendizagem por competição (GROSSBERG, 1970 a; 1972 b; 1973; 1978 a e 1987 a), teoria da ressonância adaptativa (GROSSBERG, 1976 b), implementação binária (CARPENTER & GROSSBERG, 1986a e 1987 b), implementação analógica (CARPENTER & GROSSBERG, 1987 a; 1987 c; 1987 d).

Com o passar dos anos muitos outros pesquisadores continuavam a trabalhar na teoria de sistemas neurais artificiais. Shun-Ichi Amari é um desses pesquisadores. Ele realizou estudos sobre a aprendizagem competitiva em redes e da teoria matemática das redes neurais (AMARI, 1977 e 1983). Kumihiko Fukushima também contribuiu para o avanço tecnológico das redes neurais artificiais. Em seu mais notável trabalho desenvolveu uma classe de redes neurais chamada de *neocognitron* (FUKUSHIMA & MIYAKE, 1983), aplicada ao reconhecimento de padrões visuais.

A partir da década de 70, Tuevo Kohonen realizou pesquisas em memória associativa (KOHONEN, 1984). Paralelamente outros pesquisadores como Anderson (ANDERSON et al.,

1977) e Steinbuch também o fizeram. Nesta mesma década Rumelhart & McClelland combinando seus esforços, inspirados pelo sistema de reconhecimento de voz "HEARSAY" da Universidade de Stanford, começaram seus trabalhos investigando as redes neurais artificiais pela formulação de um modelo semi-paralelo do processo mental. Posteriormente, Rumelhart & McClelland estenderam suas pesquisas para outras arquiteturas de redes neurais artificiais. Usaram o Processamento Paralelo e Distribuído (PDP) para descreverem seus trabalhos. Eles formaram um grupo de pesquisadores dedicados a estudar a estrutura mental da cognição. Os esforços do grupo de pesquisa PDP removeram a maior barreira existente para o treinamento de redes, através da introdução de novos conceitos e regras de aprendizagem.

A partir da década de 80 as redes neurais artificiais já mostravam grande parte do seu potencial. Em 1982 o interesse em redes neurais artificiais uma vez amadurecido, foi renovado. E vários eventos convergiram para fazer desta o pivô central do ano. Partiu-se então, decisivamente, para a área da aplicação prática. Robert Hecht-Nielsen, fundador de um programa de pesquisas sobre neurocomputação, construiu dois neurocomputadores dedicados ao paradigma das redes neurais artificiais, contribuindo de forma significativa em vários aspectos: engenharia, pesquisas, instrução e comercialização. Como instrutor, foi um dos primeiros a oferecer um curso que apresentava sucintamente o teorema que prevê a convergência do algoritmo backpropagation para uma larga faixa de mapeamento e como comerciante, formou sua própria companhia dedicada a por em prática aplicações reais de Sistemas Neurais Artificiais (HECHT-NIELSEN, 1987).

A partir de então, vários programas de redes neurais artificiais foram desenvolvidos. John Hopfield apresentou seu artigo sobre redes neurais artificiais à Academia Nacional de Ciências dos Estados Unidos, mostrando que as redes neurais podiam encontrar mínimos de energia e, com isso, dar soluções a complicados problemas de otimização. A partir da análise global feita por Hopfield, criou-se um grande interesse em aplicações ao paradigma das redes neurais artificiais, um novo impulso foi dado à ciência, mais especificamente à área das redes neurais artificiais. A lista das aplicações que podem ser resolvidas pelo uso das redes neurais artificiais, cresceu a partir da resolução de pequenos problemas para grandes aplicações práticas. Circuitos Integrados Neurais passaram a ser fabricados.

A história das redes neurais artificiais embora tenha tido os seus momentos de crédito e descrédito, prevaleceu a perseverança dos pesquisadores em busca de conhecimentos, indo cada vez mais longe. Podemos dizer que esta história está apenas em seu primeiro estágio de desenvolvimento. “Conhecer suficientemente bem a inteligência e ser capaz de construir o seu modelo de trabalho é certamente o mais excitante problema de desafio da raça humana, (McCORDOCK, 1979).

### 3.2- O NEURÔNIO BIOLÓGICO

O neurônio biológico é a unidade fundamental constituinte do cérebro humano. Tipicamente, um neurônio biológico é mais lento que um chip de silício; no primeiro, os eventos ocorrem na ordem de  $10^{-3}$  s, enquanto que no segundo são da ordem de  $10^{-9}$  s (SHEPERD & KOCH, 1990). O cérebro contorna a aparente lentidão dos seus neurônios através da sua disposição espacial extremamente eficiente, a qual é baseada em uma intensa interconexão entre os mesmos (SHEPERD & KOCH, 1990).

Para o estudo das redes neurais artificiais, as principais características morfológicas do neurônio biológico são o núcleo ou soma e os caminhos dos fluxos de sinais nervosos. Um típico neurônio pode ser exemplificado pela célula piramidal, comum no córtex cerebral, como mostrada na Figura 3.1, onde o corpo principal da célula, o núcleo, é o único responsável pelo processamento dos estímulos nervosos, de natureza eletroquímica, que chegam até ele. O soma é composto por uma simples unidade processadora, a qual recebe e combina sinais oriundos de muitos outros neurônios e os repassa a outros neurônios.

Esta classe de células possui dois tipos de caminhos para o fluxo da informação ou sinais nervosos, a saber:

- Dendritos
- Axônios

Os dendritos são os caminhos de entrada e o axônio é o caminho percorrido pelo sinal de saída. O axônio conecta-se aos dendritos de outros neurônios através de ligações chamadas sinapses, como destacada na Figura 3.1. A quantidade de informação transferida entre os

neurônios depende unicamente da força desta ligação e, é exatamente a força da ligação sináptica que é modificada quando o cérebro aprende alguma nova informação oriunda do seu ambiente externo.

A operação normal das sinapses ocorre por meio de dois processos:

- Pré-sináptico
- Pós-sináptico

No primeiro, o impulso elétrico transforma-se em impulso químico, enquanto que no pós-sináptico, o impulso químico transforma-se em impulso elétrico.

De um modo simplificado, o neurônio recebe os estímulos de seu mundo exterior através de seus dendritos. A comunicação de um neurônio com o mundo externo depende da função do neurônio dentro do circuito neural. Se o neurônio estiver no nível de entrada, ou seja, for uma unidade que recebe os estímulos do mundo exterior (por exemplo – neurônios da retina, no circuito da visão), a comunicação, pelo lado da entrada de estímulos (dendritos), é feita através de sua sensibilidade ao mundo exterior. Nesse caso, os neurônios que recebem os estímulos no nível de entrada, são algumas vezes denominados de “sensores nervosos”.

A comunicação interna entre neurônios é realizada através de contatos especiais denominados de *sinapses*. Estes contatos são aberturas microscópicas, que relacionam dois neurônios. O impulso move-se do primeiro neurônio iniciando um impulso no segundo neurônio. Estes impulsos propagam-se somente em uma única direção, variando na razão de 10 à 120 metros por segundo. A seqüência é a seguinte: o fim de uma fibra nervosa é estimulada até o limiar ou acima de sua sensibilidade; iniciam-se as trocas químicas e elétricas, que são os impulsos. O impulso, então, estende ao fim da fibra nervosa e pode induzir um impulso em outra célula nervosa, eventualmente resultando na ativação ou inibição fisiológica desse neurônio a depender de seu nível particular de ativação.

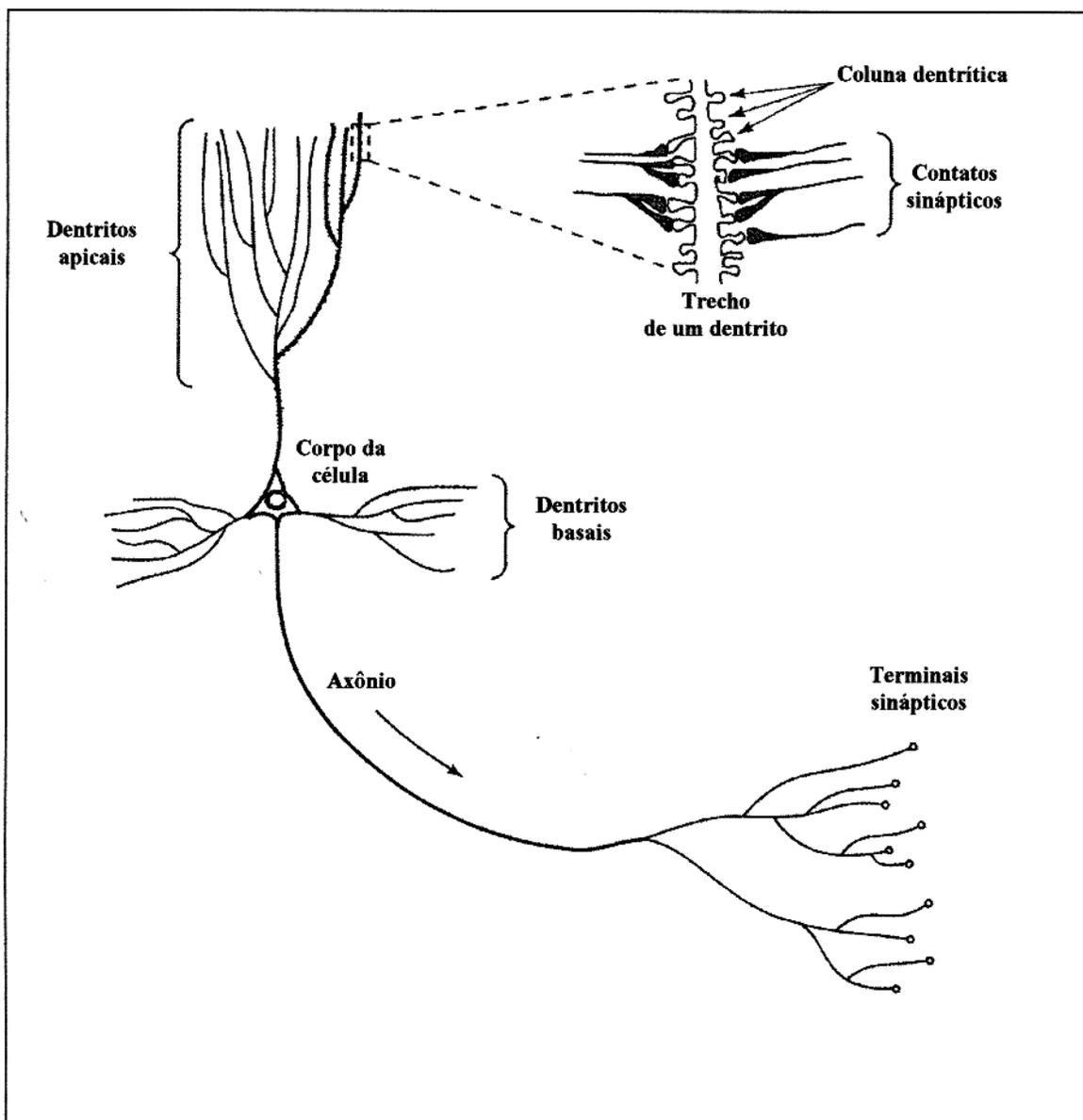


Figura 3.1 : Neurônio biológico exemplificado pela célula piramidal, comum no córtex cerebral, mostrando os caminhos do fluxo do sinal (dentritos e axônio), o corpo da célula e, no detalhe, as microcavidades das ligações sinápticas.

### 3.3- O NEURÔNIO ARTIFICIAL

O neurônio artificial é a unidade processadora fundamental para a computação das redes neurais artificiais. A Figura 3.2 mostra um modelo de um neurônio artificial típico, onde identifica-se os seus três componentes básicos:

1. Dendritos representados por um conjunto de sinapses, formado por um especificado conjunto de valores (pesos). A existência de uma conexão entre dois neurônios quaisquer  $i$  e  $j$  é caracterizada por um peso  $w_{ij}$ , onde os subíndices  $i$  e  $j$  representam o ponto de chegada e o ponto de partida, respectivamente, de um sinal que percorre a rede. Podemos distinguir dois tipos de sinapses:
  - a) Sinapse excitatória, onde  $w_{ij} > 0$
  - b) Sinapse inibitória, onde  $w_{ij} < 0$

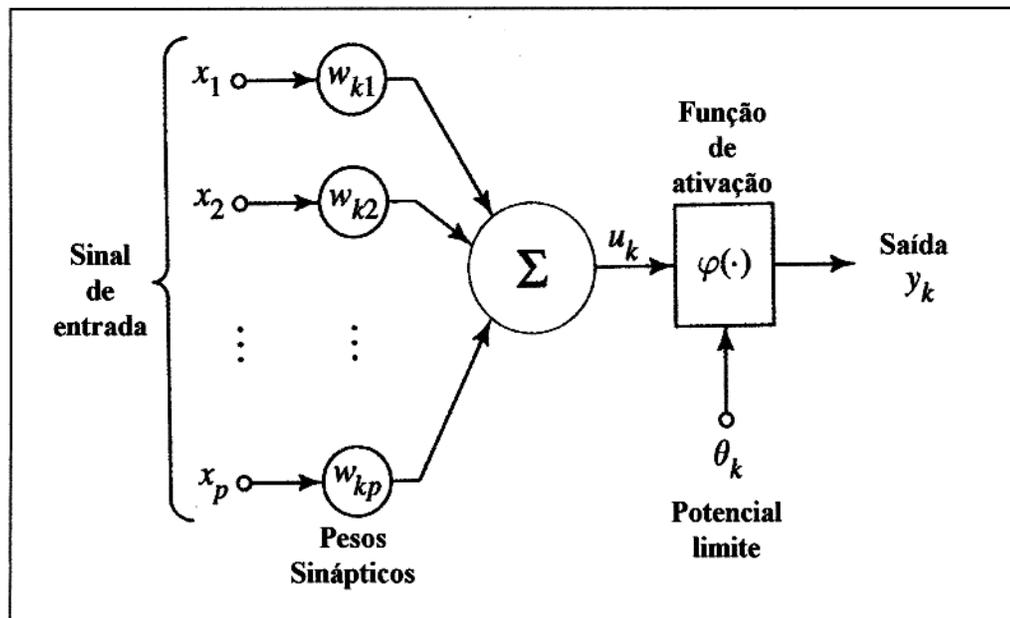


Figura 3.2: Modelo de um neurônio artificial típico, onde identifica-se os seus componentes básicos em analogia ao neurônio biológico.

2. Núcleo, constituído por dois elementos:

- 2.1. O somador, que produz uma combinação linear entre os sinais de entrada e os pesos sinápticos.

2.2. A função de ativação, responsável pela limitação da amplitude do sinal de saída do neurônio. A amplitude está comumente no intervalo fechado de  $[0,1]$  ou  $[-1,1]$ .

3. Axônio – é o sinal de saída  $y_k$ .

O modelo do neurônio artificial, inclui também um parâmetro externo  $\theta_k$ , o qual é um valor limite ou valor limiar (treshold), que deve ser excedido para que o neurônio produza um sinal de saída eficaz.

Matematicamente, pode-se descrever o estado de um neurônio  $k$ , em um instante qualquer, através do seguinte par de equações

$$u_k = \sum_{j=1}^N w_{kj} x_j \quad \text{e} \quad (3.1)$$

$$y_k = \varphi(u_k - \theta_k) \quad (3.2)$$

onde, de acordo com a Figura 3.2,  $x_1, x_2, \dots, x_N$  são os sinais de entrada de todos os  $N$  neurônios conectados ao neurônio  $k$ .  $w_{k1}, w_{k2}, \dots, w_{kN}$  são os pesos sinápticos das conexões que chegam ao neurônio  $k$ . Na equação 3.2,  $\varphi$ , representa a função de ativação, responsável pela forma de quantização do sinal de saída ( $y_k$ ) em um intervalo fechado pré-determinado. A utilização de um valor limite,  $\theta_k$ , possui o efeito de aplicar uma transformação afim (Figura 3.3) sobre  $u_k$ , como mostra a equação (3.3), para o nível de atividade  $P_k$ .

$$P_k = u_k - \theta_k \quad (3.3)$$

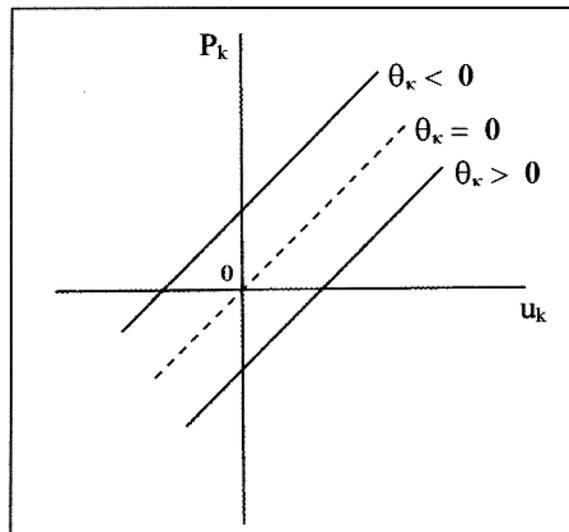


Figura 3.3: Transformação afim realizada pela introdução do termo  $\theta_k$  (potencial limite).

### 3.3.1- Funções de ativação

Uma função de ativação define a saída de um neurônio dentro de um especificado intervalo, com base em seu nível de atividade. Os três tipos fundamentais de função de ativação são:

1- Função degrau simples (treshold function), mostrada na Figura 3.4, é descrita na forma

$$\varphi(P_k) = \begin{cases} 1,0 & \text{se } P_k \geq 0 \\ 0 & \text{se } P_k < 0 \end{cases}$$

Assim, a saída do neurônio  $k$  através desta função é dada por

$$y_k = \begin{cases} 1,0 & \text{se } P_k \geq 0 \\ 0 & \text{se } P_k < 0 \end{cases}$$

Um neurônio descrito desta forma, expressa o modelo neural de McCulloh-Pitts (McCULLOH & PITTS, 1943), onde um neurônio produz uma saída efetiva somente se o seu nível de atividade for maior que zero.

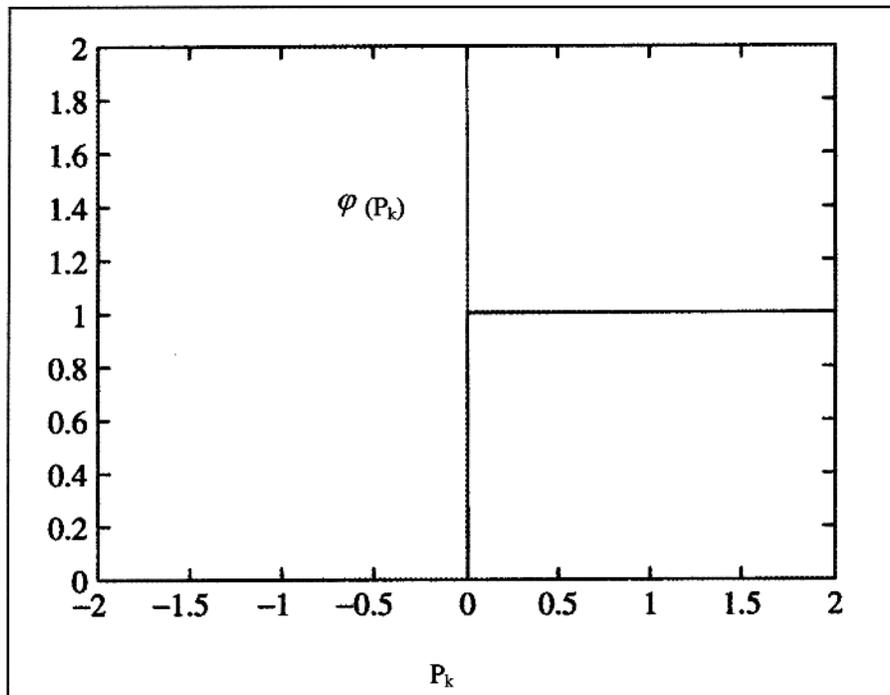


Figura 3.4: Função de ativação tipo degrau simples (treshold function).

2- Função rampa (piecewise-linear function), mostrada na Figura 3.5, é representada pela expressão

$$\varphi(P_k) = \begin{cases} 1,0 & \text{se } P_k \geq 1/2 \\ P_k & \text{se } -1/2 < P_k < 1/2 \\ 0 & \text{se } P_k \leq -1/2 \end{cases}$$

Nesta é aplicado um fator unitário de amplificação dentro da região linear. Esta forma de função de ativação pode ser vista como uma aproximação de um amplificador linear. A função rampa reduz-se à função degrau se o fator de amplificação for tomado muito grande.

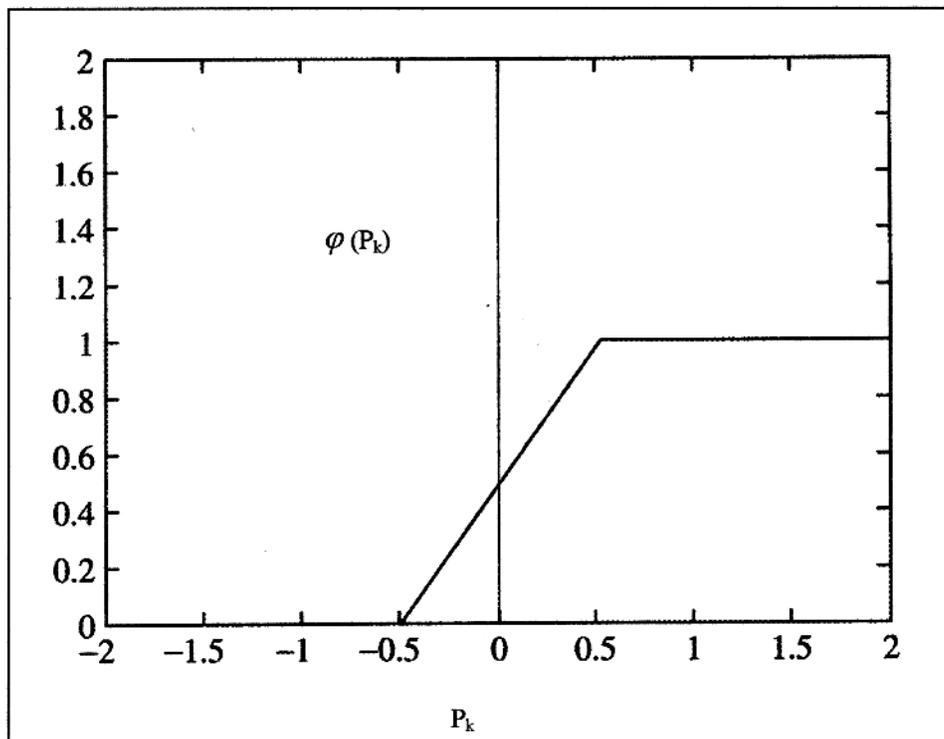


Figura 3.5: Função de ativação tipo rampa (piecewise-linear function).

3- Função sigmóide (sigmoid function), definida como uma função contínua, suave e estritamente crescente (Figura 3.6). É a função de ativação mais utilizada na construção das redes neurais artificiais, aparecendo, comumente, em duas formas:

(I) Função logística (logistic function), definida através da expressão

$$\varphi(P_k) = \frac{1}{1 + e^{(-\gamma P_k)}}$$

onde  $\gamma$  é o parâmetro de inclinação da curva. A variação deste parâmetro implica na obtenção de funções com diferentes inclinações. No limite, quando o parâmetro de inclinação tende a infinito, a função logística tenderá à função degrau simples. Enquanto que a função degrau pode assumir exclusivamente os valores 0 ou 1, a função logística pode assumir qualquer valor do intervalo fechado  $[0,1]$ .

( II ) Função tangente hiperbólica (hyperbolic tangent function), definida por

$$\varphi(P_k) = \frac{1 - e^{-\gamma P_k}}{1 + e^{-\gamma P_k}}$$

A qual produz valores de saída do neurônio no intervalo fechado  $[-1,1]$ , permitindo a existência de valores negativos, o que possui evidências neuropsicológicas de natureza experimental (EECKMAN & FREEMAN, 1986).

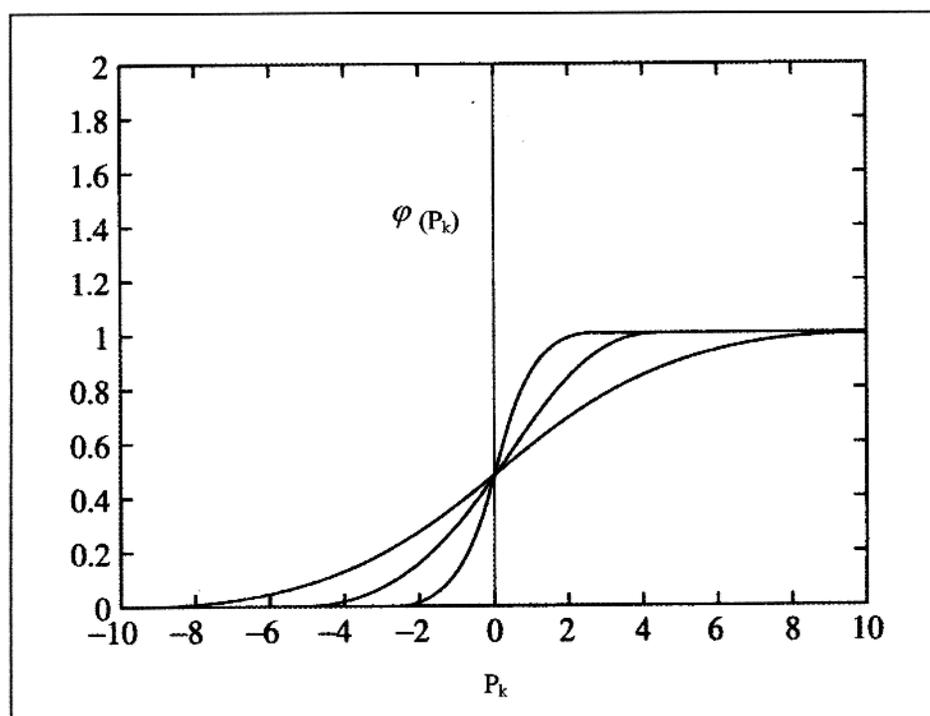


Figura 3.6: Função de ativação tipo logística, mostrando a variação da transição entre os dois estados (0 e 1), em função do parâmetro  $\gamma$ .

### 3.4- REDES NEURAIS ARTIFICIAIS

As redes neurais artificiais, também chamadas de sistemas neurais artificiais, neurocomputadores, processadores de distribuição paralela ou modelos conexionistas, são uma tentativa de imitar a estrutura e as funções do cérebro e do sistema nervoso de criaturas vivas.

As redes neurais artificiais, de uma forma geral, tentam implementar fisicamente a maneira de funcionamento do cérebro. O grande número de elementos do processamento (neurônios) presente nas redes neurais artificiais, altamente interconectados, retratam suas características de comportamento, tais como: capacidade para processar de forma paralela e eficiente a enorme quantidade de informações recebidas; habilidade de aprender com novas situações (treinamento); capacidade de generalizar (respondendo corretamente a situações novas, porém, de algum modo, relacionadas a situações anteriores).

Uma definição das redes neurais artificiais pode ser encontrada em (HECHT-NIELSEN, 1988 a):

*Uma rede neural é uma estrutura de processamento de informação paralela e distribuída que consiste de elementos de processamento (que podem possuir memória local e efetuar operações de processamento de informações locais) interconectados via canais unidirecionais chamados conexões. Cada elemento de processamento tem uma única saída que pode ser conectada a tantas outras quanto se deseje (cada uma utilizando o mesmo sinal de saída). O sinal de saída do elemento de processamento deve ser completamente local; isto é, ele deve somente depender dos valores do sinal de entrada, no momento, e dos valores armazenados na memória local do elemento de processamento.*

Uma rede neural artificial é um sistema de processamento de sinais composto de um grande número de elementos processadores simples, chamados neurônios artificiais. Estes neurônios estão interconectados por ligações chamadas conexões neuronais ou sinapses, as quais contribuem para a realização do processamento de distribuição paralela, com o intuito de resolver o problema computacional desejado.

Um dos aspectos principais da rede neural artificial é a capacidade que ela possui de adaptar-se a diversos tipos de condições ambientais através de mudanças na sua força conectiva ou estrutura.

Algumas vezes, as redes neurais artificiais são consideradas, a grosso modo, um modelo simplificado do cérebro humano. Esta visão é um tanto quanto enganosa, pois o cérebro humano ainda não foi totalmente decifrado e o seu comportamento é muito complexo. Em termos de comparação, torna-se mais sensato compararmos a capacidade de uma rede neural artificial ao sistema nervoso simples de animais primitivos, como insetos, que têm a capacidade de adaptar-se a ambientes complexos.

O estudo das redes neurais artificiais resultou do modelo computacional que imita o comportamento cerebral de um modo extremamente simplificado.

A rede neural possui dois aspectos que lembram o cérebro humano:

- A aquisição de conhecimento é realizada através de um processo de aprendizagem.
- A força das conexões interneurônios, conhecidas como pesos sinápticos, são utilizadas para o armazenamento do conhecimento.

De maneira análoga ao neurônio bioquímico, a rede neural artificial possui o seu neurônio artificial ou neurônio ou elemento processador. Um neurônio característico possui inúmeros caminhos de entrada e, normalmente, um único caminho de saída.

O funcionamento básico de um neurônio é o de promover uma combinação linear dos sinais de entrada, que posteriormente serão modificados por meio de uma função, referida como a função de ativação.

Uma rede neural artificial consiste de um grande número de neurônios que são arranjados em grupos e que recebem a designação de camadas (Figura 3.7). As camadas classificam-se em:

- Camada de entrada
- Camadas ocultas
- Camada de saída

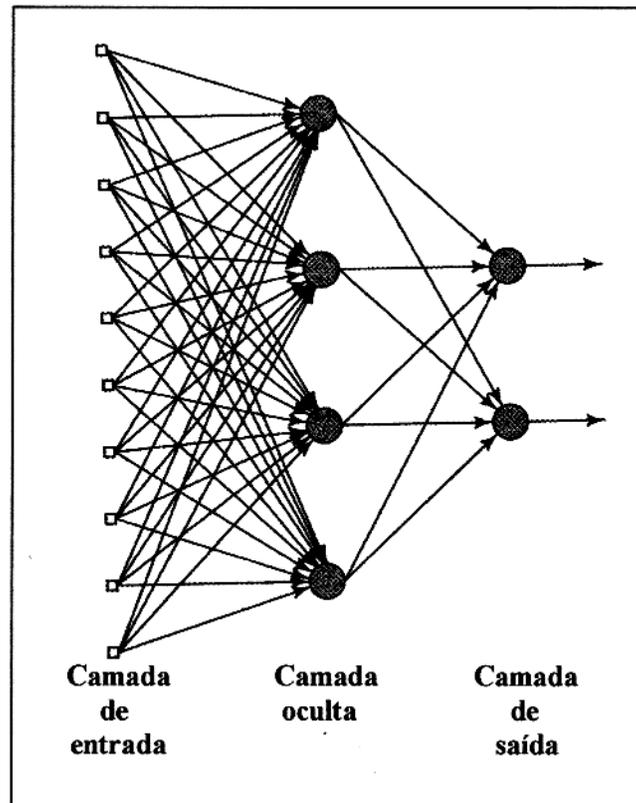


Figura 3.7: Rede neural artificial direta, com três camadas, mostrando a completa interconexão entre os neurônios de cada camada consecutiva.

As camadas da rede estão dispostas de forma ordenada, de acordo com o sentido do fluxo do sinal direto que a percorre. Possuem designação especial a primeira e a última camada. A primeira camada é chamada de *camada de entrada*, constituída por neurônios cuja única função é receber os estímulos do ambiente no qual a rede está imersa e lançá-los para o interior da rede.

A última camada recebe o nome de *camada de saída*, sendo constituída por elementos processadores e responsabilizando-se pelo retorno do sinal processado para o ambiente.

Todas as camadas dispostas entre as camadas de entrada e saída recebem a designação genérica de *camadas ocultas*, constituídas por elementos processadores, que se responsabilizam por todo o processamento que sofre o sinal no interior da rede.

No âmbito da computação com as redes neurais, é comum o estabelecimento de dois ambientes:

- Ambiente externo
- Ambiente interno

O ambiente externo é o responsável pela formação do sinal de entrada na rede e o ambiente interno, que corresponde à própria arquitetura da rede, responsabiliza-se pelo processamento do sinal de entrada e, conseqüentemente, da formação do sinal de saída que retorna ao ambiente externo.

Para a construção das redes neurais artificiais deve-se determinar como os neurônios serão interligados e definir os valores dos pesos sinápticos adequadamente. Estes determinarão como um neurônio influenciará em todos os demais a ele conectados.

O comportamento das redes neurais artificiais depende tanto dos pesos sinápticos como da função de ativação, que é a responsável pela manutenção do valor do sinal de saída dos neurônios no interior de um intervalo definido. A função de ativação pode ser específica para cada neurônio ou para cada grupo de neurônios.

Na computação de uma rede neural artificial existem duas fases principais, que são:

- Treinamento ou aprendizagem
- Operação

### **3.4.1- Treinamento ou aprendizagem**

Na fase de treinamento ou aprendizagem os pesos sinápticos são adaptados em resposta a um exemplo do conjunto de dados de treinamento. Uma característica essencial de qualquer rede neural artificial é sua regra de aprendizagem, a qual especifica como os pesos sinápticos serão adaptados em resposta a um exemplo do conjunto de treinamento da rede. O treinamento requer, sempre, a apresentação de muitos exemplos à rede neural artificial um grande número de vezes.

Temos três tipos básicos de regras de aprendizagem:

- Regra de Hebb

- Regra do Delta
- Regra da Competição

Mostraremos, a seguir, a conceituação de cada regra:

#### 3.4.1.1- Regra de Hebb:

O peso de uma conexão sináptica na entrada de um neurônio é incrementado se o estímulo e a resposta desejada, neste neurônio, são altas. Em termos biológicos, isto significa que uma conexão é fortalecida toda vez que é utilizada (HEBB, 1949).

#### 3.4.1.2- Regra do Delta :

A aprendizagem é realizada a partir de um processo contínuo de redução do erro existente entre a resposta atual de um neurônio à um estímulo e a sua correspondente resposta desejada (WIDROW & STEARNS, 1985), expresso na forma  $w_{ij}(t+1) = w_{ij}(t) + \Delta w_{ij}(t)$ .

#### 3.4.1.3- Regra da Competição:

Na qual cada neurônio compete com todos os outros até que um único vença ou produza a maior resposta à um dado estímulo (FUKUSHIMA, 1975).

A essência do treinamento de uma rede neural artificial é formular uma codificação funcional de relação estímulo-resposta. Uma rede neural é dita como sendo bem treinada, no sentido em que, se ela conhece bem o passado, poderá, então, estimar o futuro convenientemente (HAYKIN, 1994).

O treinamento pode ser visto como um problema da escolha da melhor codificação, dentro das possíveis codificações aceitáveis. Com base nisto, lançamos mão da técnica estatística da validação cruzada (cross-validation), a qual serve como guia (JANSSEN et al., 1988). Ela estabelece um procedimento padrão a ser utilizado no treinamento e na aplicação da rede neural, na forma:

1- O conjunto de dados disponível é aleatoriamente particionado em dois conjuntos,

- Conjunto de treinamento.
- Conjunto de teste.

- 2- O conjunto de treinamento é novamente particionado em dois conjuntos,
- Um, utilizado para a estimativa da melhor codificação.
  - Outro, utilizado para a validação da codificação escolhida.

Temos que escolher a melhor codificação, sobre a qual será treinada a rede e todo o seu conjunto de treinamento. A aplicação da rede treinada será então realizada sobre o conjunto de teste.

Uma característica importante da rede neural artificial é o modo como ela adquire conhecimento. No estudo das redes neurais artificiais, o conhecimento pode ser conceituado na forma (HAYKIN, 1994):

*“ O conhecimento refere-se à acumulação de informações ou modelos utilizados por uma pessoa ou máquina para interpretar, prever e, apropriadamente, responder ao ambiente exterior.”*

A memória de uma rede neural artificial, representada pela matriz de pesos sinápticos, é considerada distribuída e associativa. O termo distribuído refere-se ao fato de que a informação contida no conjunto de treinamento é armazenada por todos os neurônios da rede, que respondem de modo a formar uma memória associativa, no sentido em que uma vez que um estímulo incompleto ou deturpado (perturbado com ruído) é apresentado, ela buscará em sua memória um estímulo armazenado que melhor se ajuste ao apresentado, produzindo uma resposta que corresponda ao estímulo completo.

### 3.4.2- Operação

A fase de operação refere-se à maneira como a rede neural artificial processa os estímulos apresentados à camada de entrada e cria uma resposta na camada de saída.

O modo de operação de uma rede neural pode ser sincronizado ou não sincronizado. No modo sincronizado todos os neurônios produzem uma resposta simultânea a cada instante de tempo. No modo não sincronizado cada neurônio, em uma camada, produz o seu sinal de saída independente dos demais.

### 3.4.3- Classificação geral das redes neurais

Através do modo como os estímulos são processados, temos dois tipos de redes:

- Rede neural artificial direta (feedforward neural network).
- Rede neural artificial recorrente (recurrent neural network).

Na rede neural direta (Figura 3.8) o sinal apresentado à camada de entrada segue de forma direta passando pelas camadas ocultas até a camada de saída, camada por camada.

Nas redes neurais recorrentes (Figura 3.9) existem conexões de retorno, que ligam camadas posteriores a camadas anteriores, ocasionando a reverberação da informação entre as camadas, até que alguma convergência seja encontrada e a informação seja passada para a camada de saída.

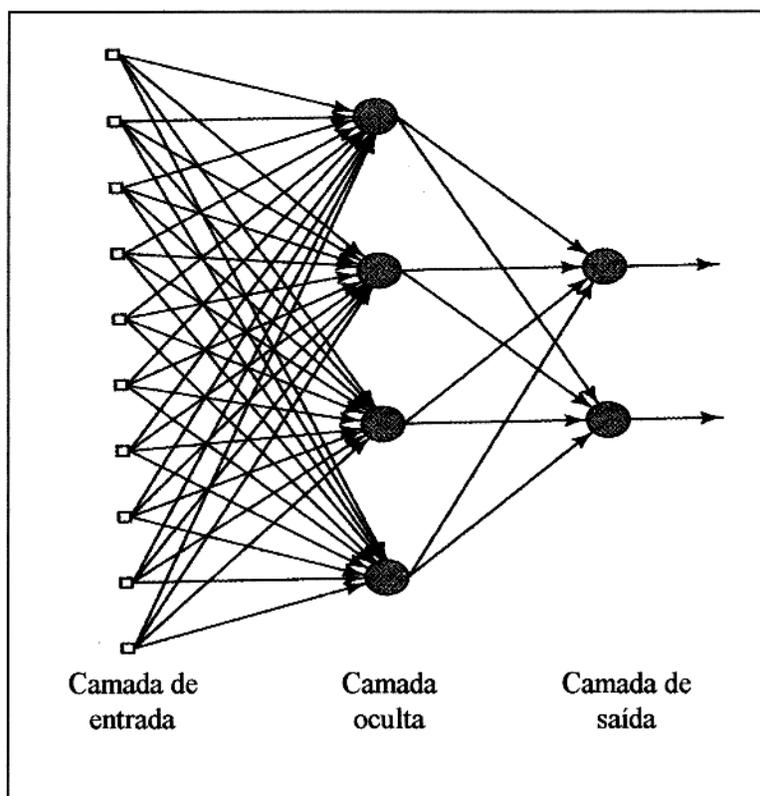


Figura 3.8: Rede neural artificial direta mostrando, através das setas apresentadas sobre as ligações sinápticas, o caminho percorrido pelo sinal, a partir da camada de entrada até a camada de saída.

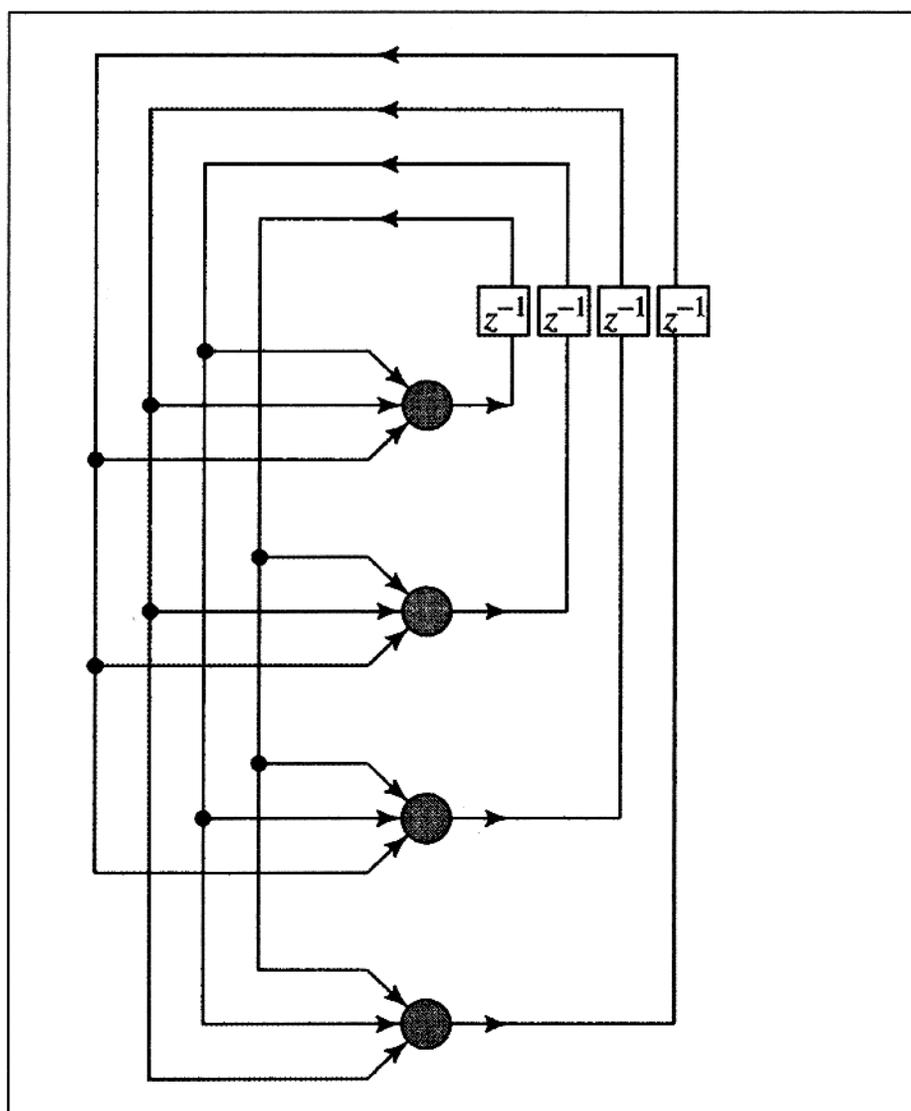


Figura 3.9: Exemplo de uma arquitetura de rede neural recorrente, mostrando as conexões entre os neurônios e a recorrência representada por  $Z^{-1}$  (atraso unitário).

### 3.4.4- Generalização

O que se espera de uma rede neural artificial treinada é que ela seja capaz de generalizar, como exemplificado pela Figura 3.10 e pela Figura 3.11, que mostram duas possibilidades distintas para o mapeamento do conjunto de estímulos.

A generalização é a principal característica de uma rede neural artificial e a responsável pela sua capacidade de solucionar problemas.

A generalização de uma rede neural artificial é influenciada por três fatores, que são:

- Tamanho do conjunto de treinamento
- A arquitetura da rede neural
- A complexidade do problema a ser tratado

Assim, percebendo que, em princípio não existe controle sobre o terceiro fator, podemos tratar a construção das redes neurais, que possibilitem uma boa generalização, sob dois aspectos (HUSH & HORNE, 1993):

- A arquitetura da rede é fixa → o problema é determinar o tamanho do conjunto de treinamento.
- O conjunto de treinamento é fixo → o problema é determinar a melhor arquitetura da rede neural.

Embora os dois aspectos sejam, em teoria, relevantes na prática; o primeiro caso é mais comumente encontrado e a determinação do tamanho do conjunto de treinamento será tratado particularmente para cada caso estudado neste trabalho.

### 3.5- REDE NEURAL DIRETA MULTICAMADAS

A arquitetura típica deste tipo de rede neural é composta por um conjunto de camadas dispostas ordenadamente em função da direção de propagação do sinal no seu interior (Figura 3.12). A primeira camada (camada de entrada) é responsável pelo recebimento das informações do meio exterior; os neurônios que a constituem não têm nenhuma finalidade computacional, sendo por isto chamados de unidades sensoriais. As camadas ocultas são constituídas por neurônios processadores responsáveis pelo processamento interno da rede. A camada de saída é formada por elementos processadores onde aparece o sinal de saída da rede.

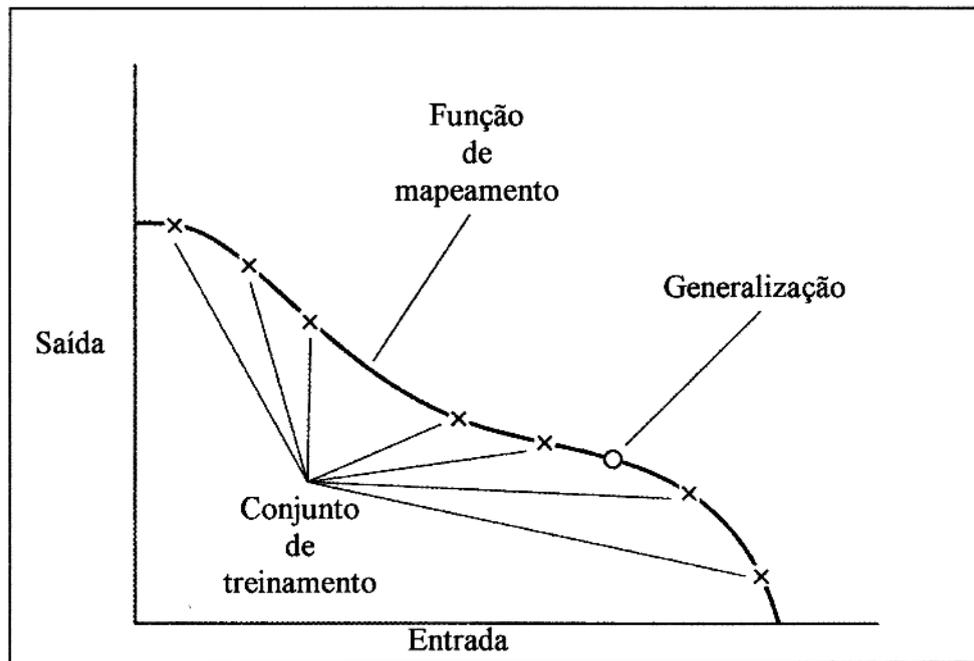


Figura 3.10: Demonstração da generalização realizada pela rede neural direta. A função de mapeamento dos estímulos é suave, produzindo uma boa generalização.

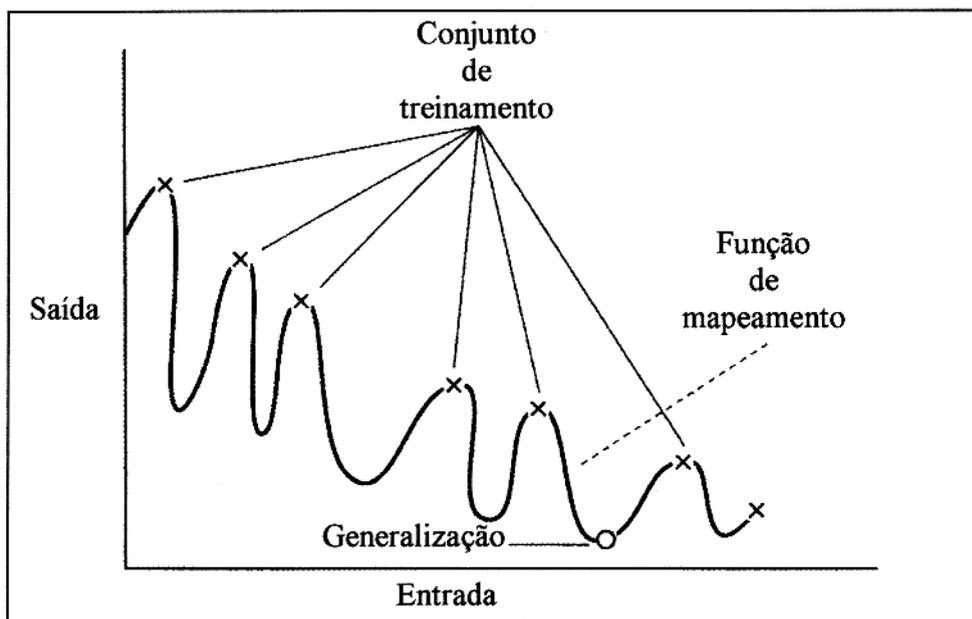


Figura 3.11: Demonstração da generalização realizada pela rede neural direta. A função de mapeamento dos estímulos não é suave, produzindo uma má generalização.

A rede neural direta de múltiplas camadas é conhecida também como perceptrons de múltiplas camadas (multilayer perceptron). Ela realiza um mapeamento não linear do espaço de estímulos (sinal de entrada) no espaço das respostas desejadas, através de regras específicas de aprendizagem, a partir da apresentação de exemplos de treinamento compostos por pares estímulo-resposta. A rede neural direta de múltiplas camadas possui as seguintes características:

- O modelo da rede inclui uma não linearidade (função de ativação) caracterizada por uma função contínua, suave e diferenciável.
- As camadas ocultas possibilitam o aprendizado de problemas complexos, extraíndo progressivamente mais características significantes dos padrões de entrada.
- A rede exibe um alto grau de conectividade, determinado por suas sinapses. Uma variação na conectividade implica em uma variação no número de sinapses ou no valor de seus respectivos pesos sinápticos.

Na rede neural direta o sinal propaga-se de forma direta; entrando pela camada de entrada, passando direto pelas camadas ocultas e retornando ao ambiente externo através da camada de saída. Assim, para este tipo de rede neural e em função do algoritmo utilizado para o treinamento da rede, podemos identificar dois tipos de sinais (PARKER, 1987), mostrados na Figura 3.13:

- Sinal direto (function signal)
- Sinal inverso (error signal)

O sinal direto é um sinal de entrada (estímulo), que se propaga do início ao fim da rede de forma direta (camada por camada) e emerge no fim da rede como sinal de saída. Ele entra na rede através da camada de entrada, passa pelas camadas ocultas e sai para o meio externo através da camada de saída.

O sinal inverso origina-se na camada de saída e propaga-se inversamente ao sinal direto (camada por camada) através da rede. Ocorre na fase de aprendizagem. A primeira camada oculta é abastecida pela camada de saída. A saída da primeira camada oculta é aplicada na segunda camada até terminarem as camadas ocultas e chegarmos na camada de entrada.

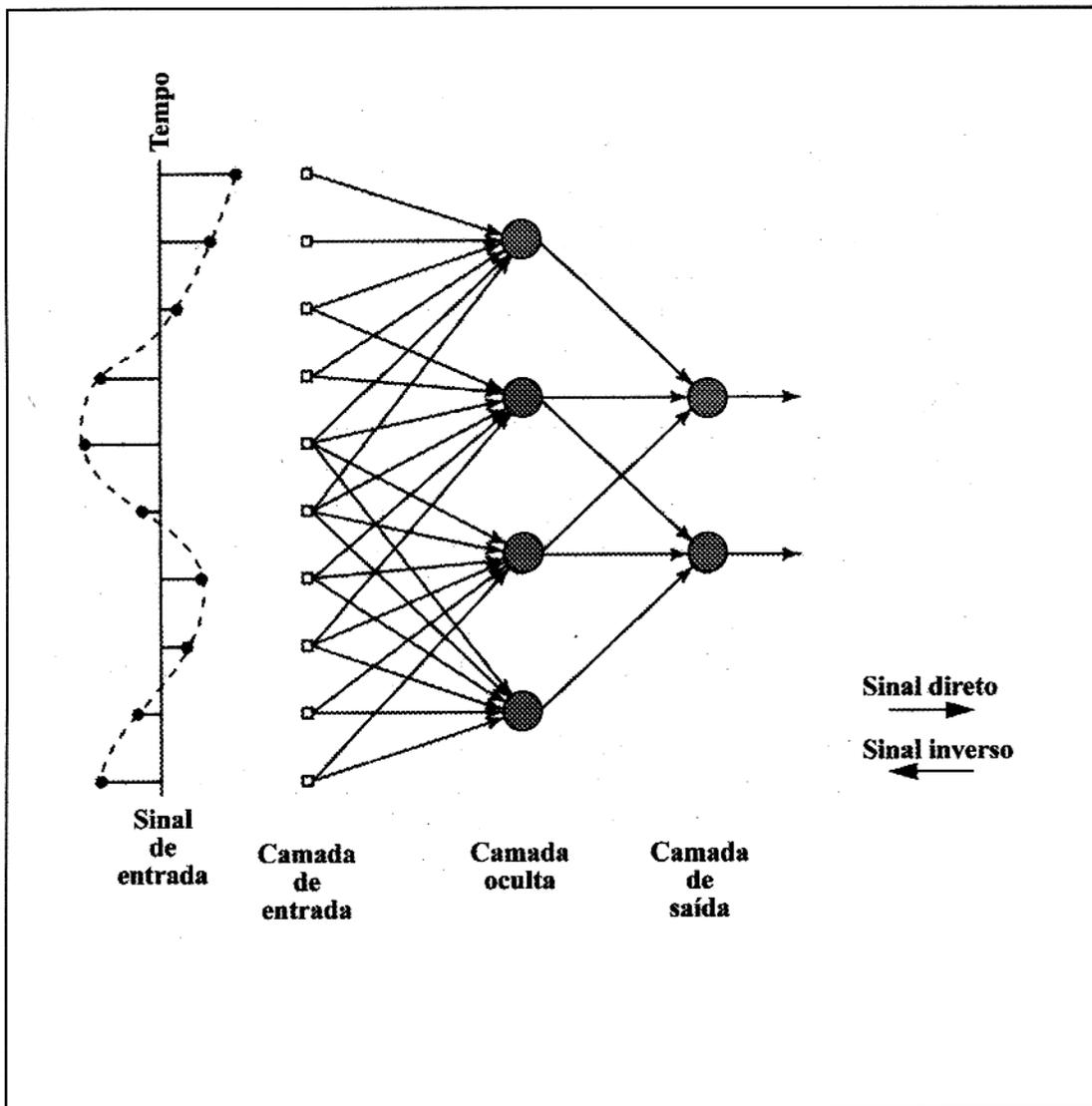


Figura 3.12: Exemplo de uma arquitetura de rede neural artificial direta, onde visualiza-se a entrada do sinal de modo discretizado, a estrutura de camadas e os dois tipos de sinal que percorrem a rede na fase de treinamento.

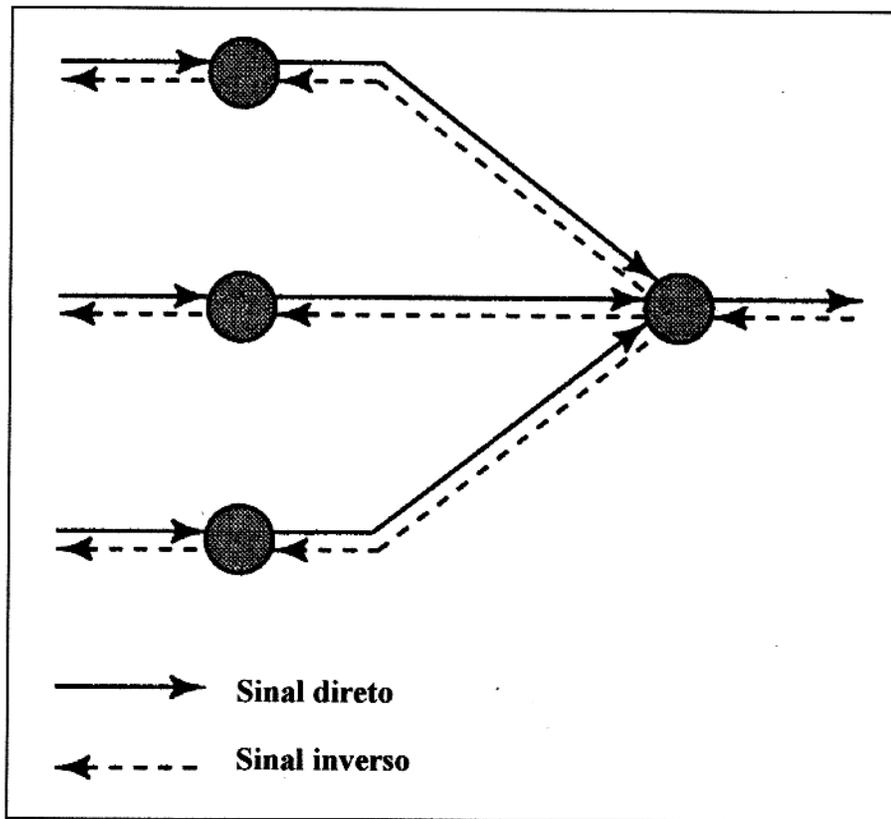


Figura 3.13: Trajetória dos sinais no interior da rede durante a fase de treinamento.

A conjugação destas características, combinadas à capacidade de aprender a partir da experiência, tornam este tipo de rede neural extremamente poderosa. No entanto, estas mesmas características são também responsáveis pelas suas deficiências. A presença de uma não-linearidade (função de ativação) distribuída por toda a rede e a utilização das camadas ocultas tornam bastante difícil a sua análise teórica, bem como a visualização do seu processo de aprendizagem. A fase de treinamento torna-se muitas vezes bastante complicada, uma vez que a pesquisa deve ser realizada sobre um espaço muito grande de possíveis funções suaves para o mapeamento estímulo-resposta desejado, além da busca de uma eficiente forma de representação dos padrões de entrada (HINTON & NOWLAN, 1987).

As pesquisas com as redes neurais diretas multicamadas iniciaram com o trabalho sobre os *perceptrons* (ROSENBLATT, 1962), concebidos a partir de um modelo da retina ocular,

constituído por um conjunto de sensores de luz distribuídos na forma de uma matriz, com o objetivo de simular a capacidade de reconhecimento de padrões do sistema visual humano. Os *perceptrons* utilizam o neurônio de Mcculloh-Pitts (McCULLOH & PITTS, 1943) e seu treinamento é realizado pela regra de Hebb. Apesar do grande interesse inicial, Minsky & Papert, em 1969, mostraram que existe um conjunto de pesos sinápticos tal que a resposta da rede coincide com a resposta desejada se, e somente se, o conjunto dos sinais de entrada são linearmente separáveis; ou seja, dado um espaço de todos os sinais de entrada, todos aqueles que pertencem a um dado conjunto ocuparão o mesmo lado do hiperplano (MINSKY & PAPERT, 1969). Outra importante contribuição foi dada por Widrow & Hoff, em 1966, com a apresentação de uma arquitetura de rede neural conhecida como *Madaline* (Multiple Adaptive Linear Neuron), construída a partir da associação de várias *Adalines* (Adaptive Linear Neurons), que são elementos processadores inspirados nos *perceptrons* e nos filtros adaptivos lineares. O seu algoritmo de treinamento utiliza a regra do Delta (WIDROW & HOFF, 1966).

As primeiras redes neurais diretas careciam, entretanto, de um melhor algoritmo de treinamento que permitisse o tratamento de problemas não lineares, o que somente foi alcançado com a publicação do livro *Parallel Distributed Processing* por Rumelhart & McClelland, em 1986, lançando o algoritmo de retro-propagação do erro (RUMELHART & McCLELLAND, 1986).

### 3.5.1- Aprendizagem por retro-propagação do erro

A base do algoritmo de retro-propagação do erro (error backpropagation) surgiu em 1974, através da tese de Ph.D. de Werbos, que propôs um algoritmo para o *cálculo do gradiente de modo reverso*. Este algoritmo foi aplicado para modelos de redes gerais podendo ser aplicado em redes neurais como um caso especial.

A retro-propagação do erro tornou-se o algoritmo de aprendizagem mais popular através de estudos desenvolvidos por Rumelhart, Hinton & William, em 1986 e, de forma independente, Parker, em 1985. A popularidade ocorreu após a publicação de um livro de dois volumes, *Processamento Paralelo Distribuído: Explorações na Microestrutura do Conhecimento*, escrito por Rumelhart & McClelland (RUMELHART & McCLELLAND, 1986).

O desenvolvimento deste algoritmo mostrou que é possível treinar eficientemente redes com camadas intermediárias, resultando no modelo de Redes Neurais Artificiais mais utilizado atualmente, as redes diretas multicamadas, treinadas com o algoritmo de retro-propagação do erro.

O algoritmo de retro-propagação do erro baseia-se na regra de aprendizagem de correção do erro ou regra do Delta. O backpropagation consiste de duas passagens do sinal através das camadas da rede neural (Figura 3.13):

- Passagem direta ou propagação
- Passagem inversa ou adaptação

Na primeira, a passagem do sinal dá-se de modo direto, não havendo alteração dos pesos sinápticos. Os dados de entrada são apresentados à rede, que os propaga através das camadas ocultas, onde cada nó calcula a soma das suas entradas e propaga a saída calculada para a camada de saída.

Na passagem inversa temos a passagem do sinal inverso, onde há o ajuste dos pesos sinápticos através da regra do Delta. Neste tipo de passagem, a saída da rede é comparada com o resultado desejado e é feito o cálculo do erro em cada nó de saída. Estes erros são transmitidos a cada nó das camadas ocultas, onde é feito o cálculo para sabermos a sua contribuição para o erro final. Com base nestes erros fazemos uma adaptação das ligações, tomando como base os seus pesos sinápticos.

Os parâmetros livres da rede são os pesos sinápticos e os potenciais limites, uma vez que o treinamento da rede trata da modificação desses valores ao longo do tempo. Os pesos sinápticos e potenciais limites, são tomados como valores aleatórios e uniformemente distribuídos no interior de um pequeno intervalo. Ao atribuímos valores iniciais a esses parâmetros livres, iniciamos, através do algoritmo de retro-propagação do erro, o processo de aprendizagem da rede (LEE et al.; 1991).

Com o início do processo de aprendizagem da rede, o algoritmo de retro-propagação do erro calcula os novos parâmetros livres através da regra do Delta, até que seja alcançado algum critério de parada pré-estabelecido. O critério de parada mais utilizado pode ser estabelecido na forma descrita a seguir (KRAMER & SANGIVANNI-VINCENTELLI, 1989).

O algoritmo de retro-propagação do erro é considerado terminado quando estabelece valores para os parâmetros livres, tal que:

- A derivada primeira da superfície de erro é próxima de zero ou de um número positivo muito pequeno.
- A soma ponderada das funções que efetuam uma medida do erro, em cada neurônio, é próxima de zero ou de um número positivo muito pequeno.

O desenvolvimento do algoritmo de retro-propagação do erro é um marco no desenvolvimento das redes neurais artificiais, uma vez que foi capaz de promover uma maneira computacional eficiente para o aprendizado das redes diretas de múltiplas camadas (HAYKIN, 1994).

### 3.5.1.1- Algoritmo de Retro-Propagação do Erro

Para mostrar a dedução do algoritmo de retro-propagação do erro, admite-se uma rede neural direta com três camadas: uma camada de entrada, uma oculta e uma de saída (Figura 3.14). O estado ou o valor do sinal direto (de saída) de um neurônio  $i$  ( $S_i(t)$ ), pertencente à camada de saída da rede no instante  $t$ , isto é, na  $t$ -ésima apresentação de um mesmo par estímulo-resposta ou na apresentação do  $t$ -ésimo par estímulo resposta, é mostrado abaixo

$$S_i(t) = \frac{1}{1 + e^{-P_i(t)}} \quad (3.4)$$

onde o parâmetro  $\theta_i$  (equação 3.2) foi retirado e incluído na expressão do potencial de ativação  $P_i$ ; incluindo-se um neurônio fictício 0, tal que, o seu estado em qualquer instante é unitário,  $S_0 = -1.0$  e o seu peso sináptico em relação ao neurônio  $i$  é  $w_{0i} = \theta_i$ .

Assim, a expressão para o potencial de ativação, considerando-se um número de  $N$  conexões com o neurônio  $i$ , fica

$$P_i = \sum_{j=0}^N w_{ij} S_j \quad (3.5)$$

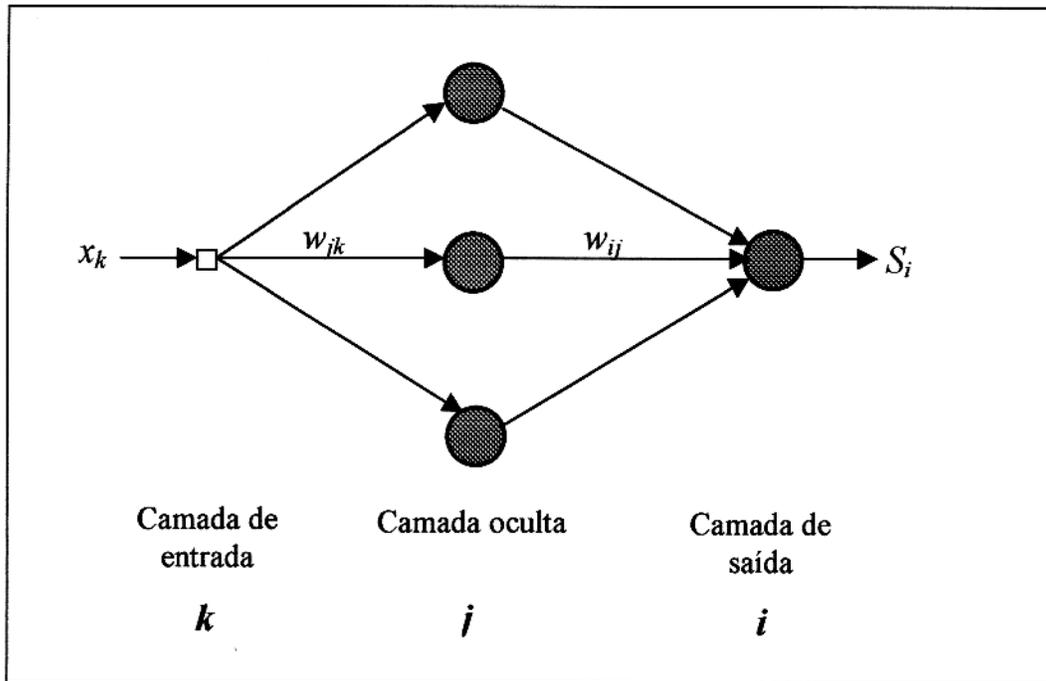


Figura 3.14: Trecho de uma rede neural direta com três camadas, onde visualiza-se o sinal de entrada ( $x_k$ ), os pesos sinápticos ( $w_{jk}$ ,  $w_{ij}$ ) e o sinal de saída da rede ( $S_i$ ).

Como o neurônio  $i$  representa um neurônio arbitrário pertencente à camada de saída da rede, o seu estado  $S_i(t)$  pode ser interpretado como um exemplo do sinal direto que se propagou pela rede desde a camada de entrada. Do mesmo modo, os sinais diretos dos neurônios situados na camada oculta e que se conectam ao neurônio  $i$  serão designados por  $S_j(t)$ .

O sinal inverso ou o sinal de erro,  $e_i$  (erro local), calculado no neurônio  $i$  da camada de saída da rede, será dado por

$$e_i(t) = D_i - S_i(t); \quad (3.6)$$

na equação acima,  $D_i$  representa a resposta desejada para o neurônio  $i$ .

A partir da equação (3.6), a função objeto ou função erro global (cost function) é definida. Esta equação é característica do algoritmo de retro-propagação do erro para a camada de saída, supondo-se, que apenas um único par estímulo-resposta seja apresentado à rede na forma

$$E(t) = \frac{1}{2} \sum_{i=1}^M e_i^2(t) \quad (3.7)$$

Interpreta-se esta equação como sendo a metade da soma quadrática de todos os erros locais dos  $M$  neurônios que compõem a camada de saída da rede no instante de tempo  $t$ . Ela é função dos pesos sinápticos  $w_{ij}$  e dos potenciais limites  $\theta_i$ , ou seja, de todos os parâmetros livres da rede.

O algoritmo de aprendizagem de retro-propagação do erro tem como objetivo básico o ajuste dos parâmetros livres da rede, de modo a promover uma minimização da função objeto  $E(t)$ . Assim, de acordo com as equações (3.4) e (3.5), torna-se necessário considerar somente o ajuste sobre os valores dos pesos sinápticos, os quais serão calculados de acordo com os respectivos erros a cada instante de tempo. O ajuste sobre o peso sináptico  $w_{ij}$ , que representa a conexão entre o neurônio  $i$  da camada de saída e o neurônio  $j$  da camada oculta, no instante  $t$ , a partir da regra do Delta, é expresso por

$$\Delta w_{ij}(t) = -\eta \frac{\partial E(t)}{\partial w_{ij}} \quad (3.8)$$

O parâmetro  $\eta$  (learning-rate parameter), na equação (3.8), é constante e responsável pela velocidade do aprendizado. O sinal negativo significa a tomada da direção decrescente do gradiente no espaço dos pesos sinápticos. Para a determinação de  $\Delta w_{ij}(t)$  torna-se necessário o cálculo da derivada parcial mostrada na equação (3.8). Para isto será utilizada a regra da cadeia do cálculo infinitesimal, obtendo

$$\frac{\partial E(t)}{\partial w_{ij}} = \frac{\partial E(t)}{\partial e_i(t)} \frac{\partial e_i(t)}{\partial S_i(t)} \frac{\partial S_i(t)}{\partial P_i(t)} \frac{\partial P_i(t)}{\partial w_{ij}(t)} \quad (3.9)$$

Desenvolvemos separadamente os termos da equação (3.9). Iniciamos com a diferenciação da equação (3.7) em relação à  $e_i(t)$

$$\frac{\partial E(t)}{\partial e_i(t)} = e_i(t) \quad (3.10)$$

Diferenciando ambos os lados da equação (3.6) em relação à  $S_i(t)$  e observando que a resposta desejada é constante durante todo o período de aprendizagem, temos que

$$\frac{\partial e_i(t)}{\partial S_i(t)} = -1 \quad (3.11)$$

Para a determinação do próximo termo, deriva-se a equação (3.4) em relação à  $P_i(t)$

$$\frac{\partial S_i(t)}{\partial P_i(t)} = \varphi'(P_i(t)) \quad (3.12)$$

O termo  $\varphi'(P_i(t))$ , na equação (3.12), representa a derivada da função de ativação em relação ao potencial de ativação do neurônio  $i$  no instante  $t$ . Diferenciando a equação (3.5) em relação ao peso sináptico  $w_{ij}(t)$ , obtemos

$$\frac{\partial P_i(t)}{\partial w_{ij}(t)} = S_j(t) \quad (3.13)$$

Para finalizar, substituímos as equações (3.10), (3.11), (3.12) e (3.13) na equação (3.9), obtendo

$$\frac{\partial E(t)}{\partial w_{ij}(t)} = -e_i(t)\varphi'(P_i(t))S_j(t) \quad (3.14)$$

Com isto, a equação (3.8) será reescrita na forma

$$\Delta w_{ij}(t) = \eta e_i(t)\varphi'(P_i(t))S_j(t) \quad (3.15)$$

Logo, para os neurônios localizados na camada de saída da rede, obtemos o sinal inverso e calculamos as modificações nos pesos sinápticos utilizando a equação (3.15).

Será considerado agora o caso dos neurônios localizados na camada oculta. É claro que não existe uma resposta desejada pré-especificada para estes neurônios e, portanto, a determinação do sinal inverso não pode mais ser realizada utilizando-se a equação (3.6). O cálculo do sinal inverso para um neurônio  $j$ , localizado em uma camada oculta, será realizado de forma recursiva, em função dos sinais inversos de todos os neurônios da camada imediatamente anterior. A primeira camada para a qual se calcula o sinal inverso é a camada de saída da rede e assim, sucessivamente, até a camada de entrada.

Para a determinação da correção  $\Delta w_{jk}(t)$  a ser aplicada ao peso sináptico  $w_{jk}(t)$  utilizaremos a equação (3.8). O peso sináptico  $w_{jk}(t)$  representa a conexão entre o neurônio  $j$

da camada oculta e o neurônio  $k$  da camada de entrada, no tempo  $t$ . Reescrevemos a equação (3.8) mostrando os novos índices, na forma

$$\Delta w_{jk} = -\eta \frac{\partial E(t)}{\partial w_{jk}} \quad (3.16)$$

A quantidade  $E(t)$ , para um neurônio  $j$  da camada oculta, será definida na forma

$$E(t) = \frac{1}{2} \sum_i^M e_i^2 = \frac{1}{2} \sum_i^M \left[ D_i - \frac{1}{1 + e^{-P_i(t)}} \right]^2 \quad (3.17)$$

Na equação (3.17), o subíndice  $i$  representa um neurônio  $i$  pertencente à camada de saída da rede.

A derivada parcial à direita na equação (3.16), a partir da regra da cadeia, pode ser escrita na forma

$$\frac{\partial E(t)}{\partial w_{jk}(t)} = \frac{\partial E(t)}{\partial S_j(t)} \frac{\partial S_j(t)}{\partial P_j(t)} \frac{\partial P_j(t)}{\partial w_{jk}(t)} \quad (3.18)$$

Discutiremos cada termo da equação (3.18), começando pelo termo  $\frac{\partial E(t)}{\partial S_j(t)}$ , para o qual aplicamos a regra da cadeia na equação (3.17), obtendo o resultado parcial

$$\frac{\partial E(t)}{\partial S_j(t)} = \sum_i^M e_i(t) \frac{\partial e_i(t)}{\partial S_j(t)} \quad (3.19)$$

Utilizando novamente a regra da cadeia, a equação (3.19) pode ser desenvolvida para

$$\frac{\partial E(t)}{\partial S_j(t)} = \sum_i^M e_i(t) \frac{\partial e_i(t)}{\partial P_i(t)} \frac{\partial P_i(t)}{\partial S_j(t)} \quad (3.20)$$

Derivando a equação (3.6) em relação ao potencial de ativação, temos

$$\frac{\partial e_i(t)}{\partial P_i(t)} = -\phi'(P_i(t)) \quad (3.21)$$

Derivando a equação (3.5) em função de  $S_j(t)$ , chega-se ao seguinte resultado

$$\frac{\partial P_i(t)}{\partial S_j(t)} = w_{ij}(t). \quad (3.22)$$

Com isto, a equação (3.20) fica escrita na forma

$$\frac{\partial E(t)}{\partial S_j(t)} = -\sum_i^M e_i(t) \varphi'(P_i(t)) w_{ij}(t). \quad (3.23)$$

A equação (3.13), escrita agora para um neurônio  $j$  da camada oculta, é dada por

$$\frac{\partial P_j(t)}{\partial w_{jk}(t)} = S_k(t) \quad (3.24)$$

e a equação (3.12) fica na forma,

$$\frac{\partial S_j(t)}{\partial P_j(t)} = \varphi'(P_j(t)).$$

A equação (3.18) pode, então, ser escrita na forma final

$$\frac{\partial E(t)}{\partial w_{kj}(t)} = -S_k(t) \varphi'(P_j(t)) \sum_i^M e_i(t) \varphi'(P_i(t)) w_{ij}(t). \quad (3.25)$$

A correção do peso sináptico  $w_{jk}(t)$  é, então, escrita na forma

$$\Delta w_{jk}(t) = \eta S_k(t) \varphi'(P_j(t)) \sum_i^M e_i(t) \varphi'(P_i(t)) w_{ij}(t). \quad (3.26)$$

## **4- IDENTIFICAÇÃO DAS PRIMEIRAS QUEBRAS COM A REDE NEURAL MULTICAMADAS**

### **4.1 – INTRODUÇÃO**

Recentes descobertas na tecnologia neurocomputacional produziram algumas novas técnicas que buscam minimizar os aspectos qualitativos na determinação das primeiras quebras e na edição de traços sísmicos, enquanto é mantida a qualidade e aceitabilidade dos dados sísmicos processados.

Neste capítulo apresentamos um método para a simulação da tomada de decisão do intérprete sísmico, na tarefa de detecção, em tempo ou profundidade, das primeiras quebras em um traço sísmico por meio de uma rede neural direta multicamadas, treinada com o algoritmo de retro-propagação do erro. Para o caso da interpretação de seções sísmicas, apresenta-se uma definição conveniente para o algoritmo e o conjunto de treinamento, tal que propicie um baixo custo computacional com a conseqüente redução do tempo de treinamento da rede.

Iniciamos este capítulo apresentando formalmente o conceito e a importância prática da expressão primeira quebra (first break). Descreveremos a abordagem do problema na forma de um problema de decisão, através da interpretação da rede neural direta funcionando como um aproximador universal de funções. Discutiremos, também, a concepção da arquitetura da rede apresentada, assim como a do conjunto de treinamento.

### **4.2 – PRIMEIRAS QUEBRAS**

Na prospecção terrestre com os métodos sísmicos, as fontes geradoras das ondas sísmicas artificiais, normalmente, são explosivas ou superficiais.

Trataremos com particular interesse a geometria do caminho do raio de uma frente de onda que se propaga do ponto de explosão diretamente para a posição dos geofones, transformando-se no primeiro sinal representativo da chegada de uma frente de onda no traço sísmico resultante.

À este particular evento no traço sísmico dá-se o nome de primeira quebra (first break), o qual representa o tempo necessário para que a onda direta percorra a distância de separação entre a fonte e o arranjo de geofones com a velocidade característica da zona de intemperismo.

Pesquisas sísmicas tridimensionais (3-D) e linhas sísmicas bidimensionais (2-D) de alta densidade são, agora, rotineiramente obtidas pela indústria com o objetivo de melhorar a resolução estrutural e estratigráfica de pequenas acumulações de hidrocarbonetos; com isto, o volume de dados sísmicos cresceu exponencialmente nas últimas três décadas, numa média de dez vezes a cada oito anos (McCORMACK et al., 1993). A tecnologia de processamento sísmico tem sido capaz de manter, mesmo com este aumento no volume de dados, a sua eficiência através das melhorias nas capacidades e nas velocidades dos computadores. Entretanto, ainda há três áreas no processamento dos dados sísmicos que não têm sua tecnologia computacional totalmente exploradas, permanecendo com suas análises na forma visual:

- Determinação das primeiras quebras
- Edição de traços sísmicos com ruído
- Análise de velocidades.

Estes três processos consomem cerca de 75% do tempo do intérprete sísmico, mas utilizam menos de 50% dos recursos computacionais disponíveis no processamento sísmico (McCORMACK et al., 1993). A razão é que estes processamentos requerem a perícia e conhecimento humano para dizer o que os eventos registrados no traço sísmico representam, pelo menos em parte, através de decisões subjetivas.

Apesar de todo o esforço técnico-científico para melhorar o tempo de processamento dos dados sísmicos, a identificação das primeiras quebras (first break picking), como mostrada na Figura 4.1, é realizada de forma manual, valendo-se do esforço visual de um intérprete sísmico.

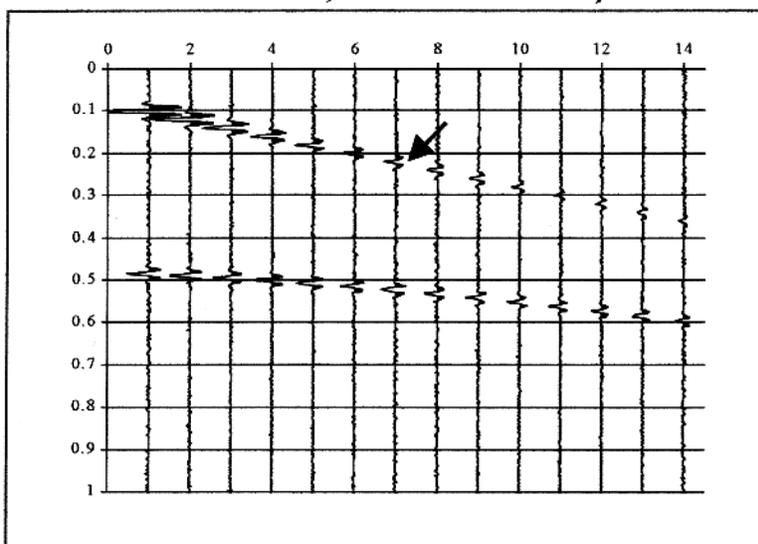


Figura 4.1: Exemplo de um registro sísmico contendo apenas dois eventos: a primeira quebra (indicada com seta) e uma reflexão na parte inferior da figura.

Hoje, ainda é bastante restrito o número de programas computacionais para a determinação automática das primeiras quebras. Esta identificação pode ser vista como um problema de reconhecimento da forma de onda representativa das primeiras quebras. Aproximações modernas tentam simular o processo de reconhecimento do intérprete. Este processo pode ser reduzido a três etapas consecutivas: Percepção - Comparação - Decisão. Trabalhos publicados recentemente lidam, principalmente, com aproximações através do algoritmo de retro-propagação do erro (error backpropagation), onde o conhecimento sobre a forma de onda representativa das primeiras quebras está armazenado em pesos que interligam os neurônios da rede.

Trabalhos publicados como o de MURAT & RUDMAN (1992) e McCORMACK et al. (1993) lidam, principalmente, com aproximações através de redes neurais que utilizam o algoritmo de retro-propagação do erro (error backpropagation). Esta aproximação é uma tentativa de simular o processo biofísico no cérebro.

Numerosas tentativas de desenvolvimento de algoritmos estatísticos ou paramétricos buscam realizar automaticamente estes problemas (ERVIN, et al., 1983; GELCHINSKY & SHTIVELMAN, 1983; NEFF & WYATT, 1986; ANDERSON & McMECHAN, 1989; SPAGNOLINI, 1991). Entretanto, estes programas apresentam um uso limitado a partir do momento que eles necessitam da robustez necessária para o processamento de dados sísmicos confiáveis, quando as condições de sinal e ruído mudam durante o curso de uma pesquisa. Além disso, a aceitação desses algoritmos varia de um intérprete para outro, dependendo se o intérprete concorda ou não com as respostas dadas pelo programa.

### 4.3 – APROXIMAÇÃO DE FUNÇÕES

Uma rede direta multicamadas treinada com o algoritmo da retro-propagação do erro pode ser vista como um processo prático para a realização de um mapeamento não linear do espaço de estímulos sobre o espaço de respostas desejadas. Para ser mais específico,  $p$  denotará o número de neurônios de entrada da rede direta multicamadas e  $q$  denotará o número de neurônios da camada de saída da rede. O relacionamento entrada-saída da rede define um mapeamento de um espaço euclidiano de entrada de dimensão  $p$  em um espaço euclidiano de saída de dimensão  $q$ , o qual é infinitamente contínuo e diferenciável.

### 4.3.1- Teorema da aproximação universal de funções

Pesquisas interessadas nas vantagens da rede direta multicamadas como dispositivo para a representação de funções contínuas arbitrárias foram, talvez, postas primeiro em foco por HECHT-NIELSEN (1987), que invocou uma versão melhorada do teorema da superposição de Kolmogorov. Então, GALLANT & WHITE (1988) mostraram que uma rede direta com uma única camada oculta apresentando a função cosseno como função de ativação, atuava como um caso especial de uma pseudo-rede neural de Fourier atuando para a aproximação de uma dada função através da Série de Fourier.

Entretanto, no conceito tradicional da rede direta multicamadas, foi CYBENKO (1988) quem demonstrou, rigorosamente, que uma camada oculta isolada é suficiente para aproximar uniformemente qualquer função contínua com suporte em um hipercubo unitário. Este trabalho foi publicado na *University of Illinois Technical Report*, em 1988, e republicado como trabalho um ano depois (CYBENKO, 1988, 1989). O treinamento aplicado à rede neural definirá um mapa do espaço euclidiano de estímulos  $p$ -dimensional, no espaço euclidiano de respostas desejadas  $q$ -dimensional, segundo o teorema (CYBENKO, 1989; HORNICK et al., 1989):

*Seja  $\varphi$  uma função não linear, contínua, limitada e monotonicamente crescente.*

*Seja  $H_p$  um hipercubo unitário,  $[0,1]^p$ ,  $p$ -dimensional. O espaço das funções contínuas no interior de  $H_p$  será denotado por  $C(H_p)$ . Então, dada qualquer função  $f \in C(H_p)$  e um  $\varepsilon > 0$ , existe um inteiro  $M$  e um conjunto de constantes reais  $\alpha_i, \theta_i$  e  $\omega_{ij}$ , onde  $i = 1, \dots, M$  e  $j = 1, \dots, p$ , tal que*

$$F(x_1, \dots, x_p) = \sum_{i=1}^M \alpha_i \varphi \left( \sum_{j=1}^p \omega_{ij} x_j - \theta_i \right) \quad (4.1)$$

*é uma aproximação da função  $f$ , ou seja*

$$|F(x_1, \dots, x_p) - f(x_1, \dots, x_p)| < \varepsilon, \quad (4.2)$$

*para todo  $x_1, \dots, x_p \in H_p$  e  $\varepsilon$  sendo um número muito pequeno.*

Este teorema mostra somente a existência, no sentido em que apresenta uma justificativa matemática para a aproximação de uma função arbitrária. A equação (4.1) é diretamente aplicada à rede neural direta se:

- A rede possui  $p$  elementos sensoriais na sua camada de entrada e uma camada oculta com  $M$  neurônios.
- Cada neurônio  $i$ , da camada oculta, possui um peso sináptico dado por  $\omega_{ij}$  e potencial limite  $\theta_i$ .
- A saída da rede é uma combinação linear das respostas dos neurônios da camada oculta e os  $\alpha_1, \dots, \alpha_M$  definem os coeficientes dessa combinação.

A equação (4.1) generaliza apenas a aproximação de funções por uma série finita, estabelecendo que uma simples camada oculta é suficiente para que uma rede neural direta produza uma aproximação da função que mapeia o conjunto de estímulos, representado por  $x_1, \dots, x_p$ , sobre o conjunto de respostas desejadas  $F(x_1), \dots, F(x_p)$ .

O teorema da aproximação universal é um teorema de existência, no sentido que ele nos dá a justificativa matemática para a aproximação de uma função contínua arbitrária como o oposto da representação exata. A equação (4.1), que é a base do teorema, simplesmente generaliza aproximações por uma série finita. De fato, o teorema nos dá que uma camada oculta isolada é suficiente para que uma rede neural obtenha uma aproximação  $\varepsilon$  uniforme para um dado conjunto de treinamento, representado pelo conjunto de entradas  $x_1, \dots, x_p$  e a saída desejada  $f(x_1, \dots, x_p)$ . Entretanto, o teorema não nos diz que uma única camada oculta isolada é a arquitetura ótima para o aprendizado.

### 4.3.2- Limites da aproximação de funções

O estudo teórico-prático que estabeleceu as propriedades da aproximação de funções contínuas arbitrárias por meio das redes neurais diretas foi realizado por BARRON (1992), assumindo uma rede com uma única camada oculta e utilizando a função sigmóide para o modelamento de cada neurônio.

Realizando o treinamento com o algoritmo da retro-propagação do erro, verificou-se que durante a fase do treinamento a rede aprende específicos pontos da função  $f$ , definidos no conjunto de treinamento. Na fase de operação, quando a rede é exposta à um novo conjunto de estímulos, ela atua na forma de um estimador de novos pontos da função  $f$ , através do estabelecimento da função  $F$  como descrito na equação (4.1).

Barron estabeleceu a independência do erro de aproximação com o número de elementos do conjunto de treinamento, o que confirma esta importante propriedade das redes neurais diretas.

### 4.3.3- Considerações práticas

O teorema da aproximação universal é importante do ponto de vista teórico, porque ele nos dá a matemática necessária para viabilizar a rede neural direta com uma única camada oculta como uma classe de soluções aproximadas. Sem tal teorema, poderíamos estar procurando por uma solução que não existe.

Entretanto, este teorema demonstra apenas a existência da rede; ele não diz como construir a rede direta multicamadas para fazer a aproximação.

Contudo, o teorema da aproximação universal tem limitado o valor prático. O teorema assume que a função contínua para ser aproximada é dada e que uma camada oculta de tamanho ilimitado está disponível para aproximação. Ambas as hipóteses foram violadas na maioria das aplicações práticas da rede direta multicamadas.

O problema com uma rede direta multicamadas utilizando uma camada oculta é que os neurônios nesse ponto tendem a interagir com cada outro globalmente. Em situações complexas, esta interação torna difícil melhorar a aproximação em um ponto sem piorá-la em algum outro ponto. Por outro lado, com duas camadas ocultas, o processo de aproximação inicia-se mais controlável. Em particular, podemos proceder como segue:

- *Características locais* são extraídas da primeira camada oculta. Especificamente, alguns neurônios da primeira camada oculta são usados para particionar o espaço da entrada dentro das regiões, e outros neurônios naquela camada aprendem as características locais caracterizando essas regiões.

- *Características globais* são extraídas da segunda camada oculta. Especificamente, um neurônio da segunda camada oculta combina a operação de saída de neurônios da primeira camada oculta numa região particular do espaço de entrada, e desse modo aprendem as características globais para aquela região.

Estes dois estágios do processo de aproximação são similares na filosofia à técnica do *spline* para o ajuste de curvas, neste sentido os efeitos dos neurônios são isolados e as aproximações em diferentes regiões do espaço de entrada podem ser ajustadas individualmente.

SONTAG (1992) deu-nos uma maior justificativa para o uso de duas camadas ocultas no contexto dos *problemas inversos*. Especificamente, o seguinte problema inverso é considerado:

*Dada uma função contínua  $f : R^p \rightarrow R^m$ , um subconjunto compacto  $C \subseteq R^m$  que está incluído na imagem de  $f$  e um  $\varepsilon > 0$ , determinam uma função contínua  $\varphi : R^m \rightarrow R^m$  tal que a seguinte condição é satisfeita*

$$\|\varphi(f(u)) - u\| < \varepsilon, \quad \text{para } u \in C.$$

Este problema chega no inverso cinemático (dinâmico), onde o estado observado  $x(n)$  de um sistema é uma função de ações correntes  $u(n)$  e o estado anterior  $x(n-1)$  de um sistema, como mostrado por

$$x(n) = f(x(n-1), u(n))$$

Assume-se que a função  $f$  é inversível, então podemos resolver  $u(n)$  como uma função de  $x(n)$  para qualquer  $x(n-1)$ . A função  $f$  representa a cinemática direta, ao passo que a função  $\varphi$  representa a cinemática inversa. Em termos práticos, a motivação é encontrar uma função  $\varphi$  que é calculada por uma rede direta multicamadas. Em geral, as funções  $\varphi$  descontínuas são necessárias para resolver o problema inverso da cinemática. (SONTAG, 1992).

#### 4.4- A REDE NEURAL DIRETA MULTICAMADAS PARA A IDENTIFICAÇÃO DAS PRIMEIRAS QUEBRAS

Nesta seção abordaremos o problema da definição ou posicionamento da primeira quebra em um traço sísmico na forma de um problema de decisão por meio de uma rede neural direta multicamadas treinada com o algoritmo da retro-propagação do erro.

Um problema de decisão pode ser interpretado como sendo a seleção de uma única amostra, sob critérios bem definidos, do interior de um conjunto finito. No caso específico da determinação das primeiras quebras, o intérprete deve selecionar no interior do conjunto finito composto por todos os eventos registrados no traço sísmico, um único evento relativo à primeira quebra, com base na sua experiência e nos critérios interpretativos dos traços sísmicos.

A simulação da tomada de decisão do intérprete, baseada na concepção de uma arquitetura de rede neural com aprendizagem supervisionada, necessita da interferência do intérprete para a definição adequada do conjunto de treinamento, representado por um conjunto de traços sísmicos escolhidos arbitrariamente da seção de tiro comum, a partir da escolha de um conjunto de traços sísmicos, nos quais ele identifica diretamente o tempo ou a profundidade da ocorrência do evento *primeiras quebras*.

As características requeridas para a simulação do sistema visual do intérprete sísmico por meio de uma rede neural são:

- O traço sísmico será a sua única fonte de informação (Figura 4.2).
- Localizar, em tempo ou profundidade, a posição de cada primeira quebra em cada traço que compõe a seção sísmica.

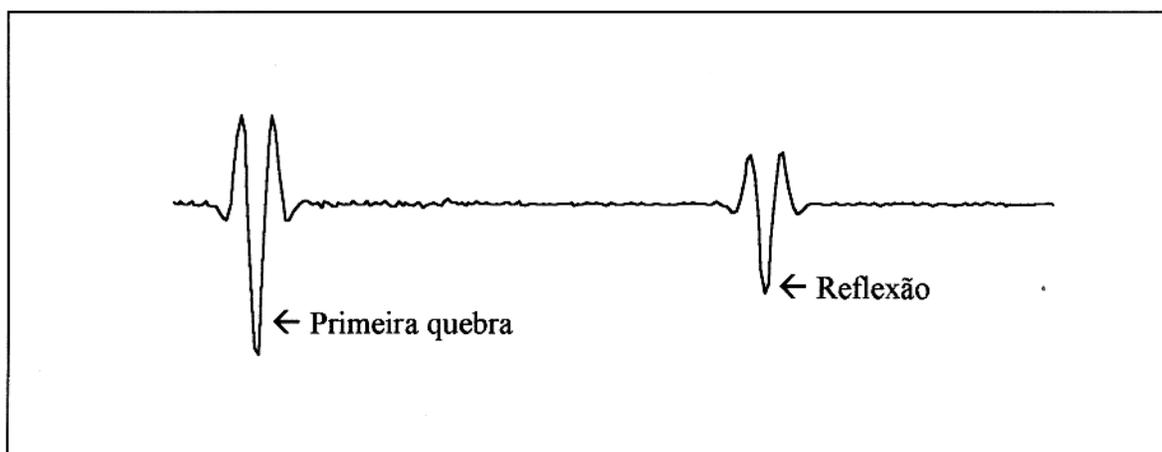


Figura 4.2: Traço sísmico sintético mostrando a primeira quebra e uma reflexão.

A exigência de que a rede, para a simulação do comportamento visual do intérprete sísmico, utilize, unicamente, o traço sísmico como única fonte de observação, deve-se basicamente a dois motivos. O primeiro visa simular exatamente o tipo de informação utilizada pelo intérprete sísmico para este trabalho, que é o traço sísmico, não cabendo portanto, para uma boa simulação, a introdução de quaisquer outras fontes de informação. O segundo motivo é de natureza puramente computacional, visando avaliar e otimizar a capacidade da arquitetura da rede na codificação e no armazenamento de informações imprecisas de natureza visual, característica do comportamento da operação do sistema visual humano.

#### **4.4.1- Arquitetura da rede**

A simulação da tomada de decisão do intérprete na determinação do posicionamento, em tempo ou profundidade, das primeiras quebras em um traço sísmico, por meio de uma rede neural direta com treinamento supervisionado, foi inspirada nas propriedades desta classe de redes na simulação do comportamento do sistema visual humano (LECUN et al., 1990; SEJNOWSKY et al., 1990; SACKINGER et al., 1992 e ANDRADE & LUTHI, 1997).

O projeto de uma arquitetura de rede neural é basicamente composto por três estágios:

- Definição do número de camadas processadoras e da forma de interconexão entre elas.
- Definição do número de neurônios em cada camada e suas respectivas funções de ativação.
- Definição do conjunto de treinamento.

##### **4.4.1.1 – Camada de entrada**

Uma forma de interpretar os dados do traço sísmico realizando o papel de estímulo (sinal de entrada) de uma rede neural direta, será entendê-lo a partir da sua própria sistemática de aquisição, na forma de uma série temporal estacionária. Esta forma de interpretação do dado sísmico simplifica, em muito, o seu relacionamento com a rede neural, no sentido em que esta poderá tratá-lo, formalmente, amostra por amostra (ANDRADE & LUTHI, 1997). Desta forma, a variável tempo deixa de ser interveniente no processo. Essa interpretação do dado sísmico induziu à concepção de uma camada de entrada para a rede neural composta por um único neurônio.

A principal característica deste tipo de interpretação do dado sísmico, associado à definição do número de neurônios da camada de entrada da rede, é a de possibilitar o

processamento de traços sísmicos com diferentes números de amostras. Este fato é de caráter geral e implica na possibilidade da utilização de uma mesma arquitetura de rede neural para o processamento de diferentes linhas sísmicas, as quais certamente possuem seções sísmicas com diferentes números de amostras ao longo do tempo. Além disso, a existência de um único neurônio na camada de entrada permite uma boa flexibilidade na escolha do tamanho do conjunto de treinamento.

#### 4.4.1.2 – Camada de saída

A definição da camada de saída da rede parte da interpretação do problema de simulação do comportamento visual do intérprete sísmico para a determinação das primeiras quebras na forma de um problema de decisão, isto é, frente ao conjunto de eventos que formam o traço sísmico (reflexões, múltiplas, primeira quebra, ruídos, etc...) o intérprete decidirá com base na sua experiência se um determinado evento representa ou não o registro de chegada da onda direta. Desta forma, tem-se um problema de decisão do tipo binário, com duas únicas alternativas: sim, para o caso de primeira quebra e não, para todos os outros eventos.

Essa interpretação do problema condicionará, no projeto da arquitetura da rede neural, na necessidade da existência de um único neurônio na sua camada de saída, o que, em outras palavras, seria a associação direta da resposta binária característica deste neurônio à solução do problema por um especialista. Neste caso, onde a rede detectar a primeira quebra, o neurônio da camada de saída estará ativado, produzindo um sinal de saída ideal igual a um. Para todos os outros eventos, este neurônio estará desativado e o seu sinal de saída será nulo.

#### 4.4.1.3 - Camada oculta

Para completar o projeto da arquitetura da rede neural, resta a definição do número de neurônios e, conseqüentemente, o número de camadas ocultas. Esta definição depende, fundamentalmente, da forma de interpretação do processamento envolvido na rede neural e do tipo de problema com o qual a rede operará. Neste trabalho interpreta-se a rede neural, treinada com o algoritmo da retro-propagação do erro, na forma de um aproximador universal de funções. Conforme visto anteriormente, sob este ponto de vista, necessita-se de apenas uma camada oculta. Por outro lado, a definição do número de neurônios nesta camada dependerá exclusivamente da complexidade dos padrões de treinamento.

O número de neurônios de cada camada oculta é variável e depende da complexidade do trecho do traço sísmico contendo a primeira quebra. O critério adotado neste trabalho, com base nos experimentos realizados, foi o de estabelecer o número inicial de neurônios da camada oculta da rede igual ao número de amostras do conjunto de treinamento. Essa forma possibilitou um tempo de treinamento adequado, com um bom desempenho, na fase de operação para identificação das primeiras quebras.

#### 4.4.1.4 - Síntese da arquitetura da rede

A arquitetura da rede neural, como mostrada na Figura 4.3, pode ser resumida na forma:

1. A rede neural é composta por três camadas.
2. A camada de entrada é constituída por apenas um neurônio.
3. A camada oculta é composta por um número variável de neurônios, apresentando como função de ativação a função logística ou

$$\varphi(P_j) = \frac{1}{1 + e^{-P_j}}$$

onde  $P_j$  representa o potencial de ativação do neurônio  $j$  da camada oculta.

4. A camada de saída possui apenas um neurônio que possui uma função linear como função de ativação.

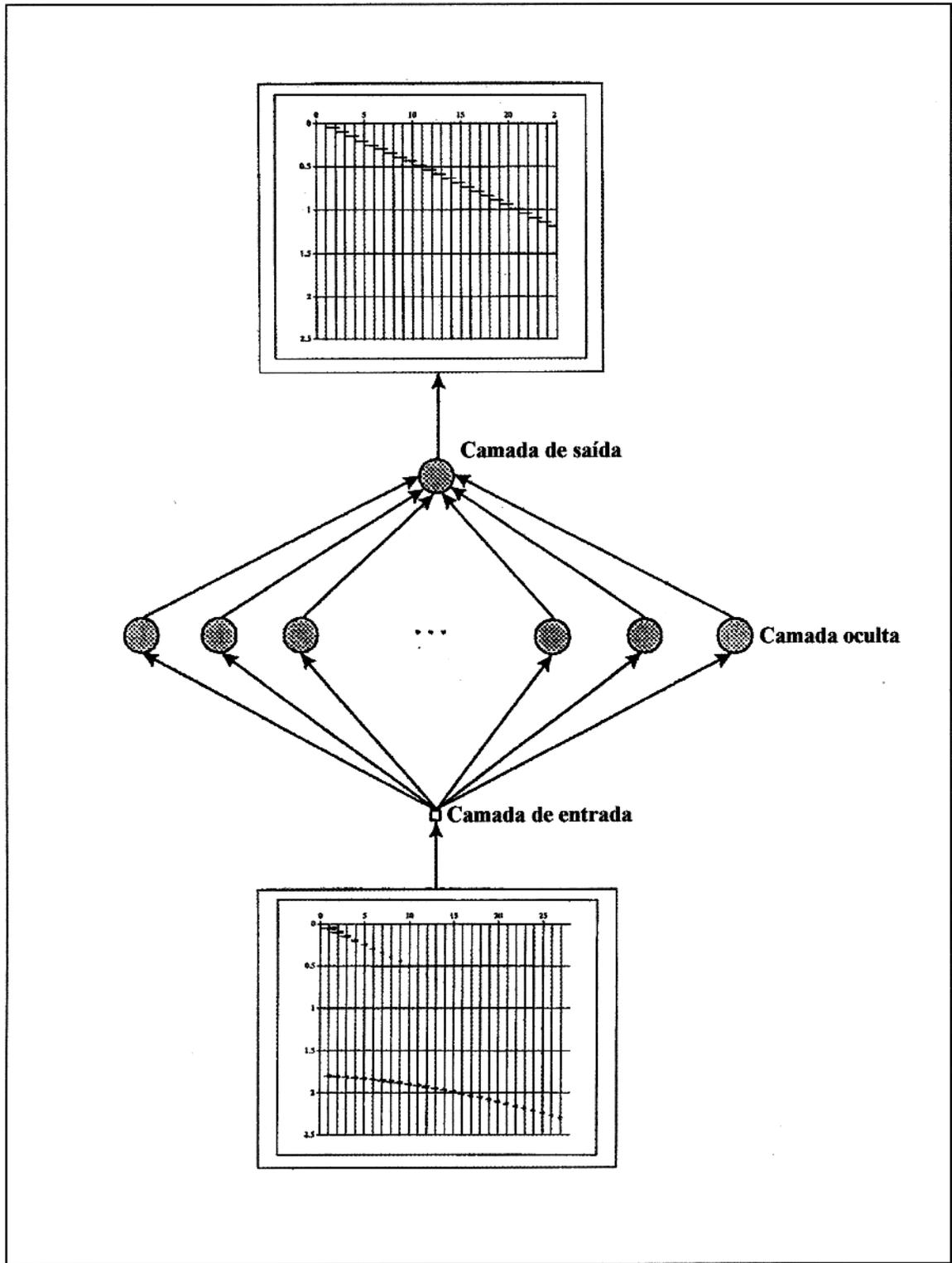


Figura 4.3: Arquitetura da rede neural para a identificação das primeiras quebras, mostrando também o dado de entrada (seção sísmica) e o dado de saída (seção sísmica contendo somente as primeiras quebras na forma de spikes).

## 4.5 – CONJUNTO DE TREINAMENTO

Conceitualmente, o conjunto de treinamento é o responsável direto pelo treinamento supervisionado da rede, nele estão contidas todas as informações referentes ao ambiente no qual a rede neural artificial irá operar. O conjunto de treinamento é composto por dois subconjuntos

- Conjunto de estímulos: é composto por todas as informações do ambiente.
- Conjunto de respostas desejadas: explicita a codificação através da qual a rede neural interpreta a informação no ambiente externo.

Inúmeros testes foram realizados, até que chegamos à uma sistematização considerada razoável para o conjunto de treinamento, de tal forma que a rede adquiriu o conhecimento necessário do seu ambiente externo. Assim sendo, podemos definir o conjunto de treinamento para a simulação da tomada de decisão do intérprete na definição das primeiras quebras no traço sísmico, como mostrada na Figura 4.4, com as seguintes características:

- O conjunto de estímulos é formado por um conjunto de trechos dos traços sísmicos, contendo primeiras quebras e seqüências de não primeiras quebras. As primeiras quebras características que formarão o conjunto de estímulos serão escolhidas a critério do intérprete.
- O conjunto de respostas desejadas é criado na forma de um conjunto binário, no qual o valor 1 (um) indica a presença de uma primeira quebra e o valor 0 (zero), para o caso contrário. A definição da posição da primeira quebra no vetor de resposta desejada, pode ser realizada de acordo com o conhecimento empírico do intérprete sísmico.

A utilização deste conjunto de treinamento sobre a arquitetura da rede neural direta, mostrou uma redução expressiva do tempo de treinamento, com a utilização do algoritmo original da retro-propagação do erro. No entanto, a utilização de uma variante desse algoritmo, como mostraremos a seguir, mostrou-se computacionalmente mais eficiente.

## 4.6 – ALGORITMO DE TREINAMENTO

A performance do algoritmo de retro-propagação do erro, na sua formulação tradicional, como visto anteriormente, mostra em muitos casos práticos uma convergência lenta (JACOBS, 1988). A utilização do algoritmo convencional de retro-propagação do erro no treinamento das redes estimuladas com dados do traço sísmico, apresentou um tempo de convergência ainda

extremamente longo. Para contornar esta situação, utiliza-se a associação de duas metodologias para promover a aceleração do processo de treinamento da rede neural.

#### 4.6.1- Termo de momento e adaptação do parâmetro de aprendizagem

A convergência do algoritmo de retro-propagação do erro é determinada pela equação abaixo

$$\omega_{ij}(t+1) = \omega_{ij}(t) + \Delta\omega_{ij}(t) \quad (4.3)$$

O ajuste do peso sináptico  $\Delta\omega_{ij}(t)$ , como apresentado anteriormente, é dado pela equação

$$\Delta\omega_{ij}(t) = \eta e_i(t) \varphi'(P_i) S_j(t) \quad (4.4)$$

O parâmetro de aprendizagem  $\eta$ , que governa o tempo de treinamento, deveria ser escolhido suficientemente pequeno de modo a promover a minimização da função erro  $E$ . Entretanto, a adoção de pequenos valores para  $\eta$  implica em uma convergência extremamente lenta. No caso contrário, para altos valores de  $\eta$  obtemos a redução do tempo de convergência. Porém, em contra partida, não teremos garantia da minimização da função erro.

A aceleração do processo de convergência e garantia de minimização da função erro, podem ser obtidos com a introdução do termo de momento  $\alpha$  na equação (4.3) ou como segue

$$\Delta\omega_{ij}(t) = \eta e_i(t) \varphi'(S_i) S_j(t) + \alpha \Delta\omega_{ij}(t-1) \quad (4.5)$$

O termo de momento reduz o tempo de convergência, mantendo a estabilidade do processo, evitando oscilações parasitas (RUMELHART & McCLELLAND, 1986; CICHOCKI & UNBEHAU, 1995). Intuitivamente, se a mudança no valor do peso sináptico anterior é alta, adiciona-se uma fração desse valor ao novo peso, acelerando o processo. Tipicamente, o termo de momento deve ser tomado tal que  $0 \leq \alpha \leq 1$ . Nos experimentos realizados, adota-se  $\alpha = 0.85$ .

Outra modificação do algoritmo tradicional de retro-propagação do erro, que causa uma drástica redução do tempo de processamento, é a possibilidade da variação do parâmetro de aprendizagem  $\eta$ , descrito matematicamente na forma

$$\eta(t) = \begin{cases} a\eta(t-1) & \text{se } E(\bar{\omega}(t)) \leq E(\bar{\omega}(t-1)) \\ b\eta(t-1) & \text{se } E(\bar{\omega}(t)) \geq \rho E(\bar{\omega}(t-1)) \\ \eta(t-1) & , \text{ outros} \end{cases} \quad (4.6)$$

É realizado então, um incremento no valor do parâmetro de aprendizagem de forma empírica, multiplicando-se  $\eta$  por  $a=1.05$ , se a função erro ( $E$ ) a ser minimizada é decrescente, isto é, o novo valor calculado do sinal erro é inferior ao anterior. Quando a função erro for crescente, promove-se uma redução do parâmetro de aprendizagem através da sua multiplicação por  $b=0.7$ , isto é, o valor atual do sinal de erro é  $\rho$  ( $\rho=1.04$ ) vezes maior que o valor anterior. Neste, o valor atual do peso sináptico e do sinal de erro serão descartados, o parâmetro de treinamento atualizado e um novo valor do peso sináptico calculado.

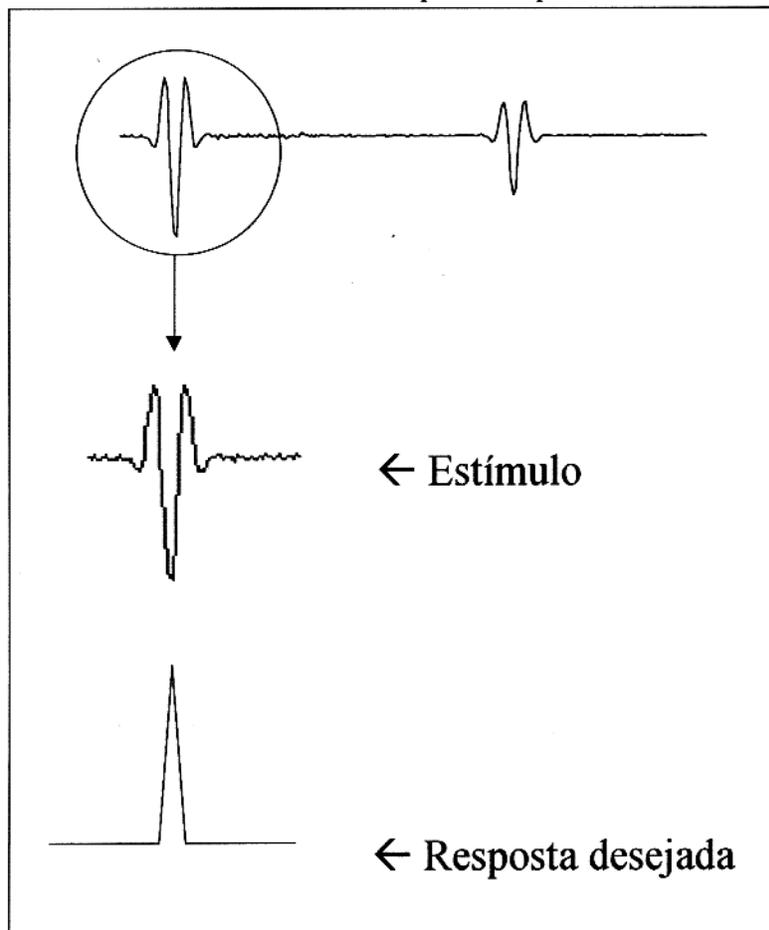


Figura 4.4: Exemplo de um conjunto de treinamento utilizado para a identificação automática das primeiras quebras; onde identifica-se por um círculo, no traço sísmico, a forma de onda representativa das primeiras quebras, a qual formará o estímulo. Na parte inferior da figura, mostra-se a resposta desejada da rede neural.

## 5- RESULTADOS

### 5.1- INTRODUÇÃO

Ao longo deste capítulo, discute-se os vários testes aplicados sobre a rede neural direta apresentada no capítulo anterior, de forma a avaliar o seu desempenho na identificação das primeiras quebras nos traços sísmicos; descritos aqui através de dados sintéticos obtidos com a teoria do raio. Essa avaliação dar-se-á sobre modelos geofísicos constituídos por camadas planas, horizontais ou inclinadas.

### 5.2- DADOS SÍSMICOS SINTÉTICOS

Os dados sísmicos sintéticos utilizados para as avaliações da rede neural direta, foram obtidos a partir do programa *seis88* (CERVENY & PSENCIK, 1988); que foi projetado para promover os cálculos computacionais relativos ao traçamento dos raios das ondas sísmicas, geradas por uma fonte pontual em subsuperfície e que chegam ao sistema de geofones distribuídos, regular ou irregularmente, ao longo da superfície terrestre. Os correspondentes tempos de trânsito são calculados automaticamente; bem como as amplitudes e os deslocamentos de fase, considerando-se os efeitos das pequenas absorções. Todas essas quantidades são registradas e podem ser utilizadas para o cálculo e a geração de sismogramas sintéticos.

Através do programa *seis88*, as ondas diretas, as ondas refletidas,  $P$  e  $S$ , e as reflexões múltiplas de qualquer tipo, podem ser geradas, incluindo-se as ondas convertidas no ponto de reflexão.

O modelo geofísico é bidimensional e lateralmente não homogêneo, podendo apresentar interfaces curvas. A fonte pode estar situada em qualquer ponto do meio, com exceção de camadas cuja espessura seja zero, isto é, não pode situar-se entre duas interfaces coincidentes.

### 5.2.1- Modelo geofísico

O modelo geofísico é assumido como bidimensional e lateralmente não homogêneo, constituído por camadas sobrepostas com interfaces planas, horizontais ou inclinadas. Os pontos de tiro podem estar situados em uma das extremidades ou no centro do modelo.

Conforme mostrado esquematicamente na Figura 5.1,  $EE_i$  representa a espessura da camada  $i$  à esquerda no modelo e  $ED_i$  representa a espessura da camada  $i$  à direita no modelo.  $PT$  representa o ponto de tiro,  $V_i$  a velocidade da camada  $i$  e  $D$  a extensão lateral do modelo.

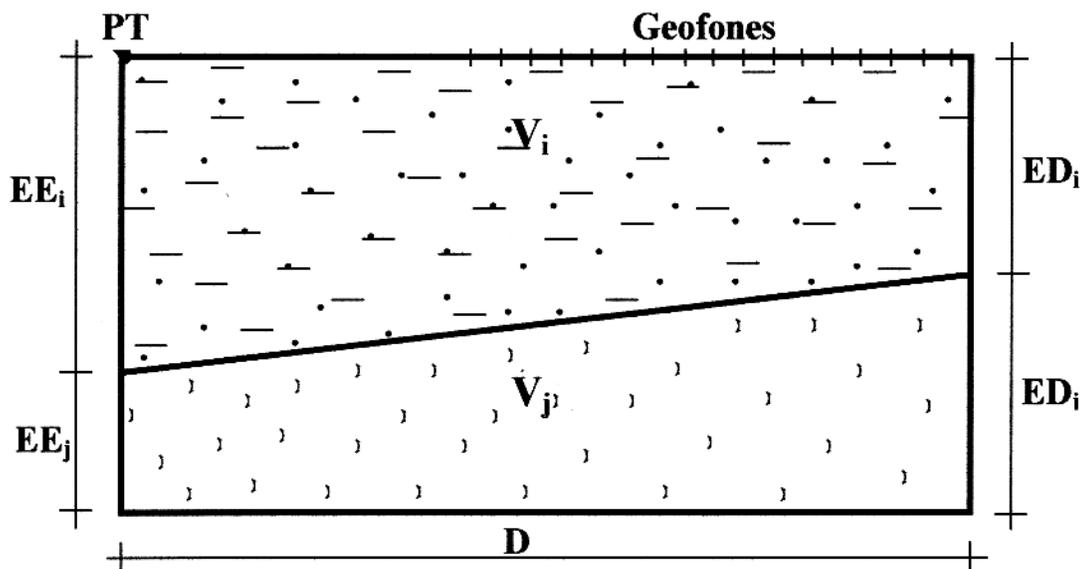


Figura 5.1: Exemplo do modelo geofísico utilizado para a geração dos sismogramas sintéticos.

### 5.3- TREINAMENTO DA REDE NEURAL DIRETA

A aplicação da técnica das redes neurais artificiais para a simulação dos processos visuais humanos, como no caso da identificação das primeiras quebras, pode ser dividida em dois estágios:

- 1- Definição da arquitetura da rede neural artificial
- 2- Definição do conjunto de treinamento

Inúmeros testes foram realizados, até que chegamos à uma sistematização considerada razoável para o conjunto de treinamento, de tal forma que a rede adquiriu o conhecimento

necessário do seu ambiente externo. Assim sendo, podemos definir o conjunto de treinamento para a simulação da tomada de decisão do intérprete na identificação das primeiras quebras no traço sísmico, com as seguintes características:

- O *conjunto de estímulos* é formado por um conjunto de trechos dos traços sísmicos, contendo as primeiras quebras e seqüências de não primeiras quebras. As primeiras quebras características, que formarão o conjunto de estímulos, serão escolhidas a critério do intérprete.
- O *conjunto de respostas desejadas* é criado na forma de um conjunto binário, no qual o valor 1 (um) indica a presença de uma primeira quebra e o valor 0 (zero), para o caso contrário. A definição da posição da primeira quebra no vetor de resposta desejada, deve ser realizada de acordo com o conhecimento específico do intérprete.

A Figura 5.2 ilustra quatro instantes do treinamento da rede neural direta, mostrando, em azul, o conjunto de estímulos, em verde, o conjunto de respostas desejadas, e, em vermelho, a resposta da rede neural naquele determinado instante.

Este conjunto de treinamento, com as devidas alterações no conjunto de estímulos, foi utilizado para o treinamento da rede em todos os casos que mostraremos a seguir.

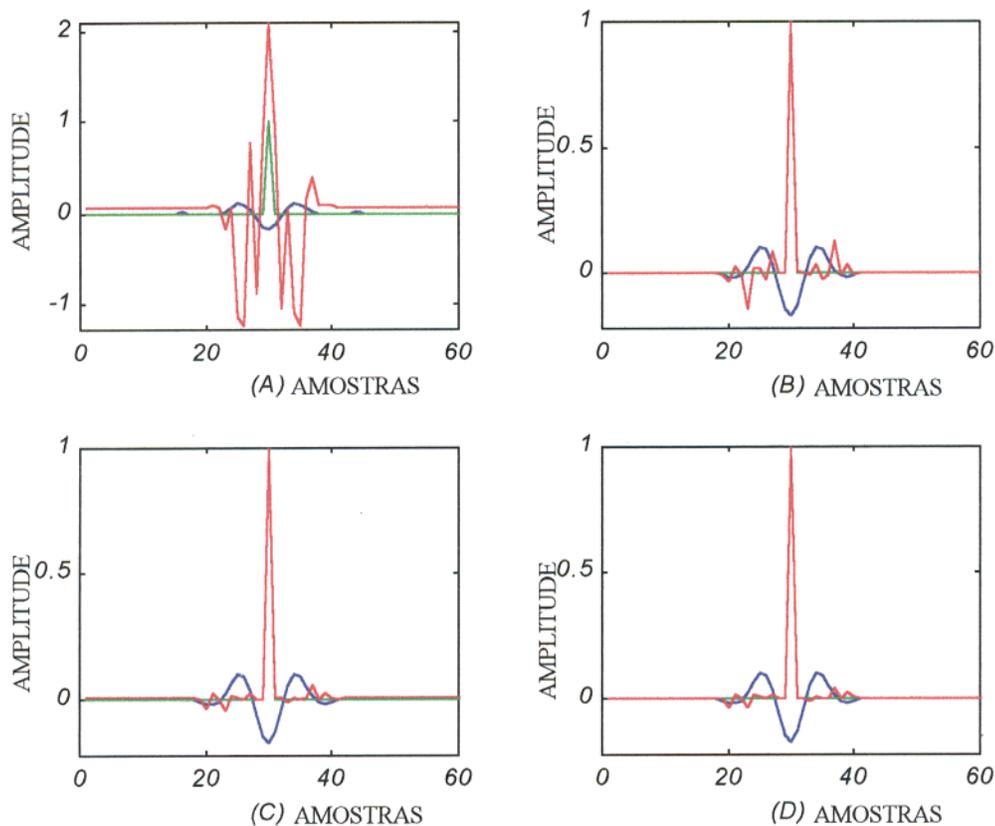


Figura 5.2: Exemplos de quatro instantes do treinamento da rede neural direta, mostrando, em azul, o conjunto de estímulos; em verde, o conjunto de respostas desejadas; e, em vermelho, a resposta da rede neural em cada instante do treinamento.

## 5.4- AVALIAÇÃO DA REDE NEURAL DIRETA

Nesta seção, avalia-se a operação da rede neural direta para a identificação automática das primeiras quebras nos traços sísmicos, em função do critério da incorporação da informação fornecida à rede na fase de treinamento.

### 5.4.1- Avaliação geral da operação da rede

A efetividade da automação da identificação das primeiras quebras será realizada diretamente sobre o caso multicanal. Na Figura 5.3, mostra-se uma seção sísmica sintética, na forma de um registro de tiro, com 11 traços sísmicos, contendo apenas dois eventos relevantes: as primeiras quebras, na parte superior da figura, e as respostas de um refletor sísmico. O conjunto de aprendizagem e as fases do processo de treinamento estão mostrados na Figura 5.2. Nos casos simples, o conjunto de estímulos pode ser constituído pela informação de uma única primeira quebra. Na Figura 5.4, mostra-se o resultado da operação da rede neural direta treinada, na forma de um registro de tiro, semelhante ao da Figura 5.3, com apenas um evento, que são as primeiras quebras, automaticamente identificadas, representadas por meio de *spikes*, que identificam os tempos de chegada da onda direta. Pode-se melhor observar que a rede neural direta foi capaz de, convenientemente, incorporar as informações da fase de treinamento e simular o comportamento visual do intérprete através da verificação da Tabela 5.1, que mostra na sua primeira coluna, o número do traço sísmico referente à Figura 5.3; na segunda coluna, os tempos de chegada das primeiras quebras, retirados do modelamento numérico de cada traço, e em sua terceira coluna, os tempos de chegada das primeiras quebras automaticamente identificadas pela rede neural direta.

Uma segunda avaliação, antes de se aplicar a rede neural direta sobre casos mais complexos, refere-se à verificação da sua acuidade em discernir entre dois eventos próximos e semelhantes, qual refere-se exatamente à primeira quebra. Na Figura 5.5, tem-se a mesma seção sísmica vista na Figura 5.3, agora contaminada por eventos incoerentes e de natureza qualquer, localizados nas proximidades das primeiras quebras em alguns dos traços. O comportamento operacional da rede neural direta submetida ao treinamento, como mostrado na Figura 5.2, pode ser avaliado pela observação da Figura 5.6, que mostra as primeiras quebras identificadas por meio de *spikes*.

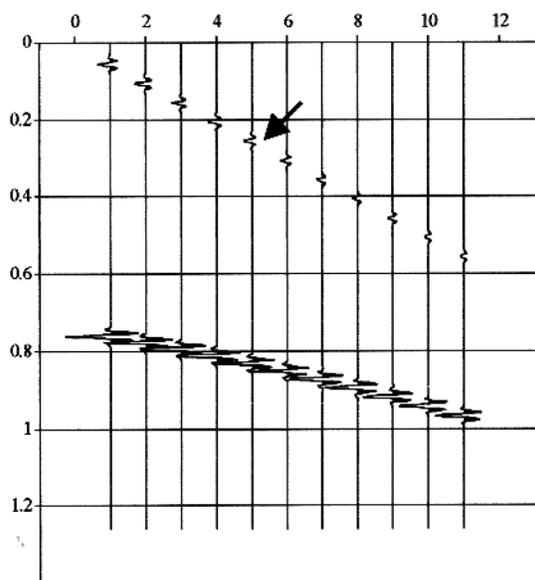


Figura 5.3: Registro de tiro sem ruído. A seta indica o traço sísmico utilizado para o treinamento da rede.

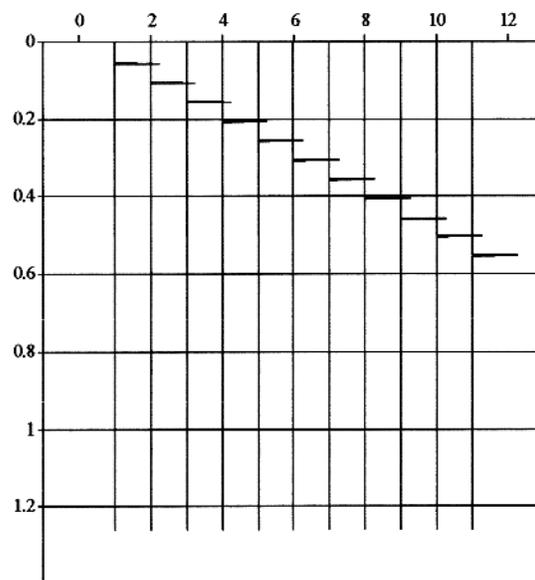


Figura 5.4: Identificação das primeiras quebras através de spikes.

Tabela 5.1: Tabela para a avaliação da precisão apresentada pela rede neural direta na identificação das primeiras quebras.

Número do traço sísmico	Primeira quebra verdadeira (s)	Primeira quebra identificada (s)
1	0.058	0.058
2	0.108	0.108
3	0.158	0.156
4	0.210	0.210
5	0.260	0.260
6	0.308	0.306
7	0.360	0.356
8	0.410	0.406
9	0.460	0.460
10	0.508	0.504
11	0.560	0.556

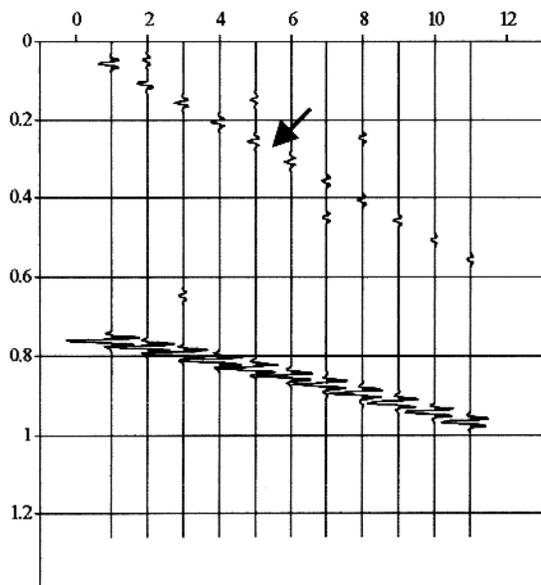


Figura 5.5: Registro de tiro com ruído. A seta indica o traço que foi utilizado para o conjunto de treinamento.

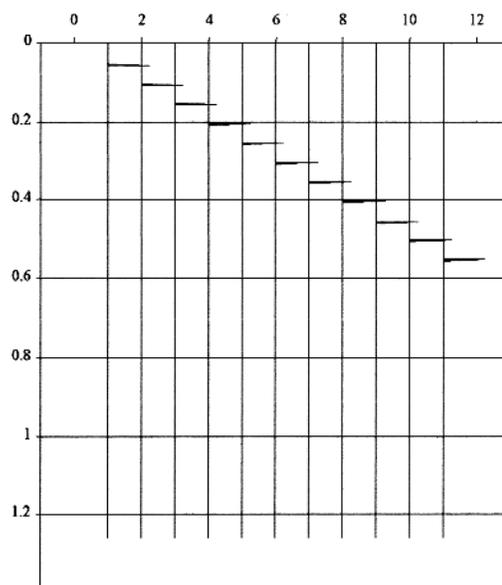


Figura 5.6: Identificação das primeiras quebras através de spikes.

## 5.5 AVALIAÇÃO COM SEÇÕES SÍSMICAS

Uma vez demonstrado a eficiência da rede neural direta na identificação das primeiras quebras, passa-se a avaliar a sua operacionalidade para o processamento de seções sísmicas com um maior número de traços e contaminadas por ruído branco.

Neste tópico será mostrado, para cada seção sísmica, o modelo geofísico utilizado para a sua geração no programa *seis88*, o sismograma sintético obtido e a resposta da rede na identificação das primeiras quebras.

A avaliação da operacionalidade da rede é realizada com seções sísmicas sintéticas, onde cada traço foi contaminado por ruído, do tipo série branca, caracterizado, no tempo, por uma sequência de valores aleatórios com distribuição uniforme e média zero. A amplitude máxima do ruído presente em cada traço é assumida como uma fração do maior valor da amplitude do sinal na seção sísmica.

### 5.5.1- Sismograma com um refletor

O modelo geofísico (Figura 5.7) utilizado para a obtenção das seções sísmicas sintéticas, mostrados nas Figuras 5.8 e 5.10, possui uma única interface. As velocidades da primeira e segunda camadas são iguais à 0,3 km/s e 3,0 km/s, respectivamente. A espessura da primeira camada é de 0,6 km e a da segunda igual à 1,25 km, sendo o comprimento lateral do modelo de 3,0 km. A fonte está situada na origem, à uma profundidade de 0,015 km; utiliza-se 25 geofones com espaçamento de 0,05 km e o offset é de 0,25 km.

Na seção sísmica mostrada na Figura 5.10, temos o sinal contaminado com ruído na forma de uma série branca, com amplitude máxima igual à um quinto da máxima amplitude registrada na seção. Na Figura 5.11, pode-se avaliar a resposta da rede neural direta para a identificação das primeiras quebras em dados sísmicos ruidosos.

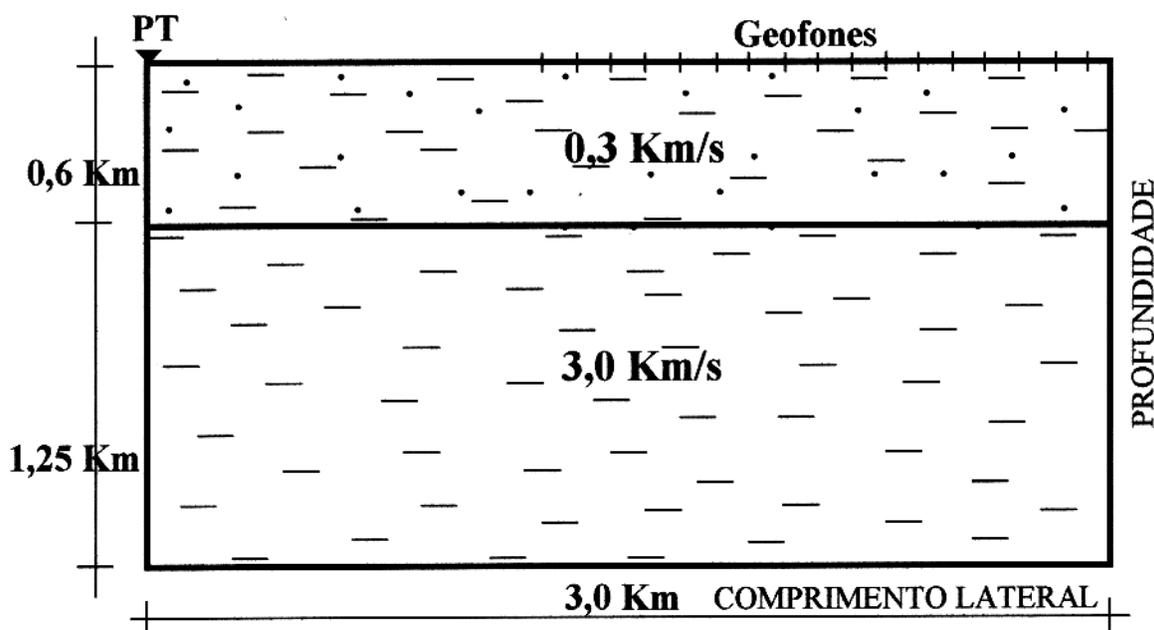


Figura 5.7: Modelo geofísico utilizado para a geração dos sismogramas sintéticos com um refletor.

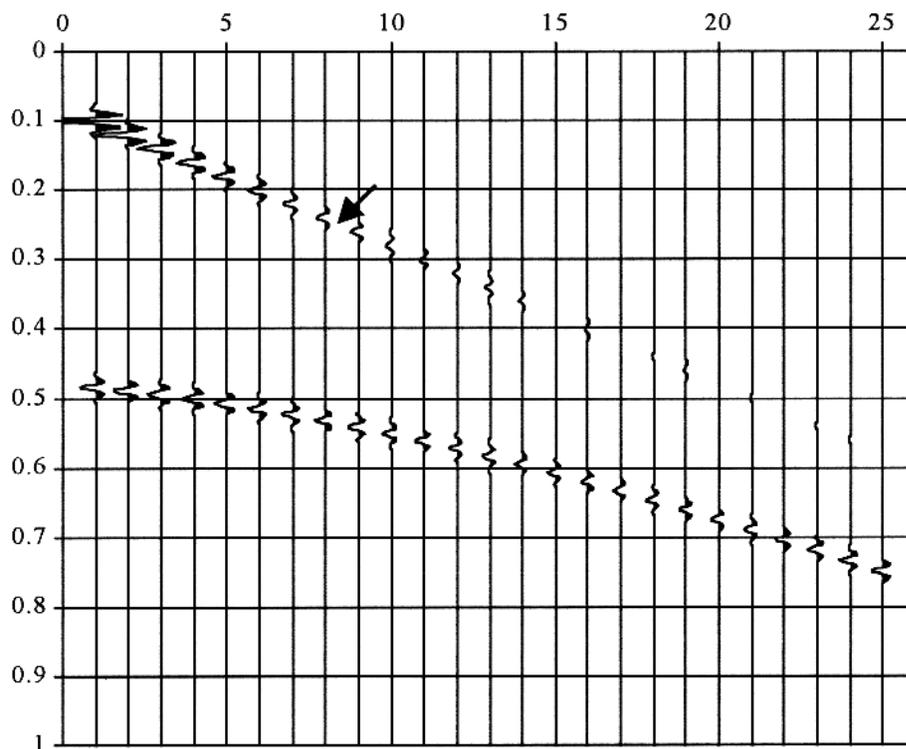


Figura 5.8: Sismograma sintético, com um refletor, sem ruído.

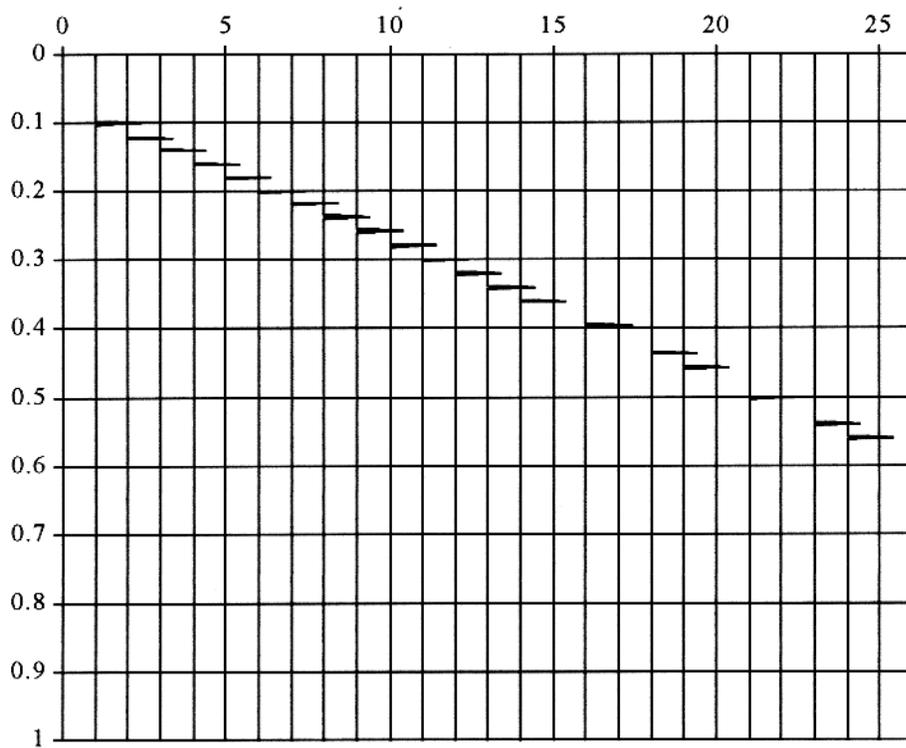


Figura 5.9: Primeiras quebras identificadas pela rede neural no sismograma sintético, com um refletor, sem ruído.

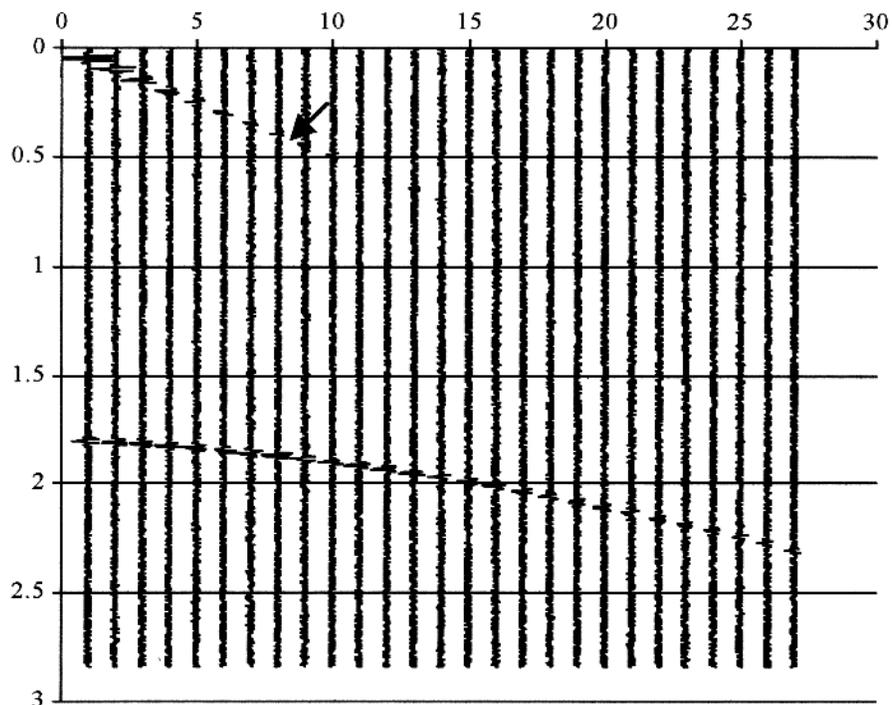


Figura 5.10: Sismograma sintético, com um refletor, com ruído.

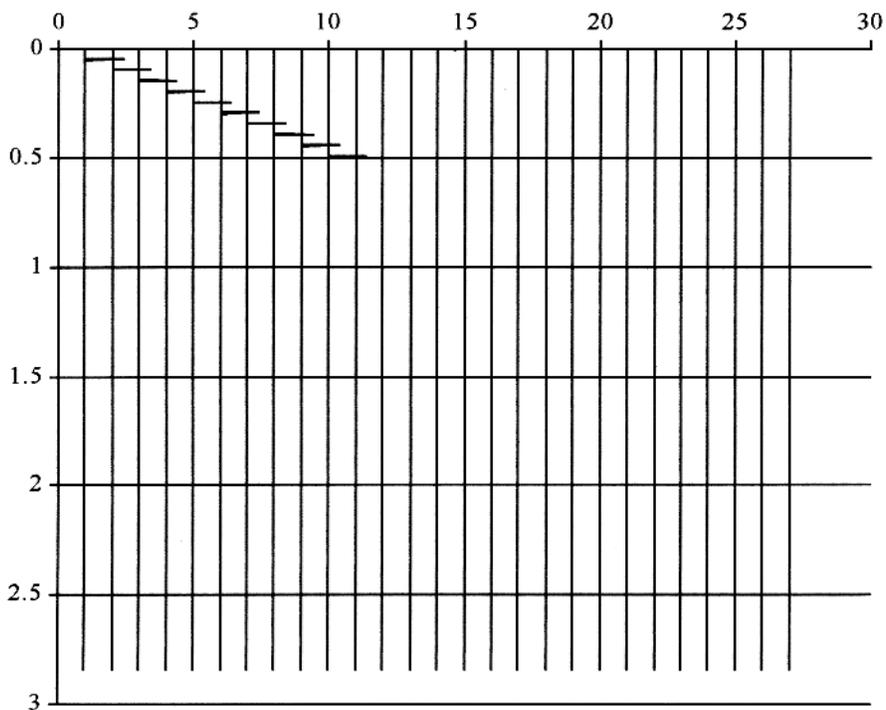


Figura 5.11: Primeiras quebras identificadas pela rede neural no sismograma sintético, com um refletor, com ruído.

### 5.5.2- Sismograma com dois refletores

O modelo geofísico (Figura 5.12) utilizado para a obtenção das seções sísmicas sintéticas mostradas nas Figuras 5.13 e 5.15, possui duas interfaces. As velocidades da primeira, segunda e terceira camadas são iguais à 0,7 km/s, 2,8 km/s e 5,0 km/s, respectivamente. As espessuras das camadas são, do topo para a base, de 0,8 km, 2,0 km e 3,0 km, sendo o comprimento lateral do modelo de 3,0 km. A fonte está situada na origem, à uma profundidade de 0,015 km; utiliza-se 25 geofones com espaçamento de 0,05 km e o offset é de 0,25 km.

Na seção sísmica mostrada na Figura 5.15, temos o sinal contaminado com ruído na forma de uma série branca, com amplitude máxima igual à um décimo da máxima amplitude registrada na seção. Na Figura 5.16 pode-se avaliar a resposta da rede neural direta para a identificação das primeiras quebras em dados sísmicos ruidosos.

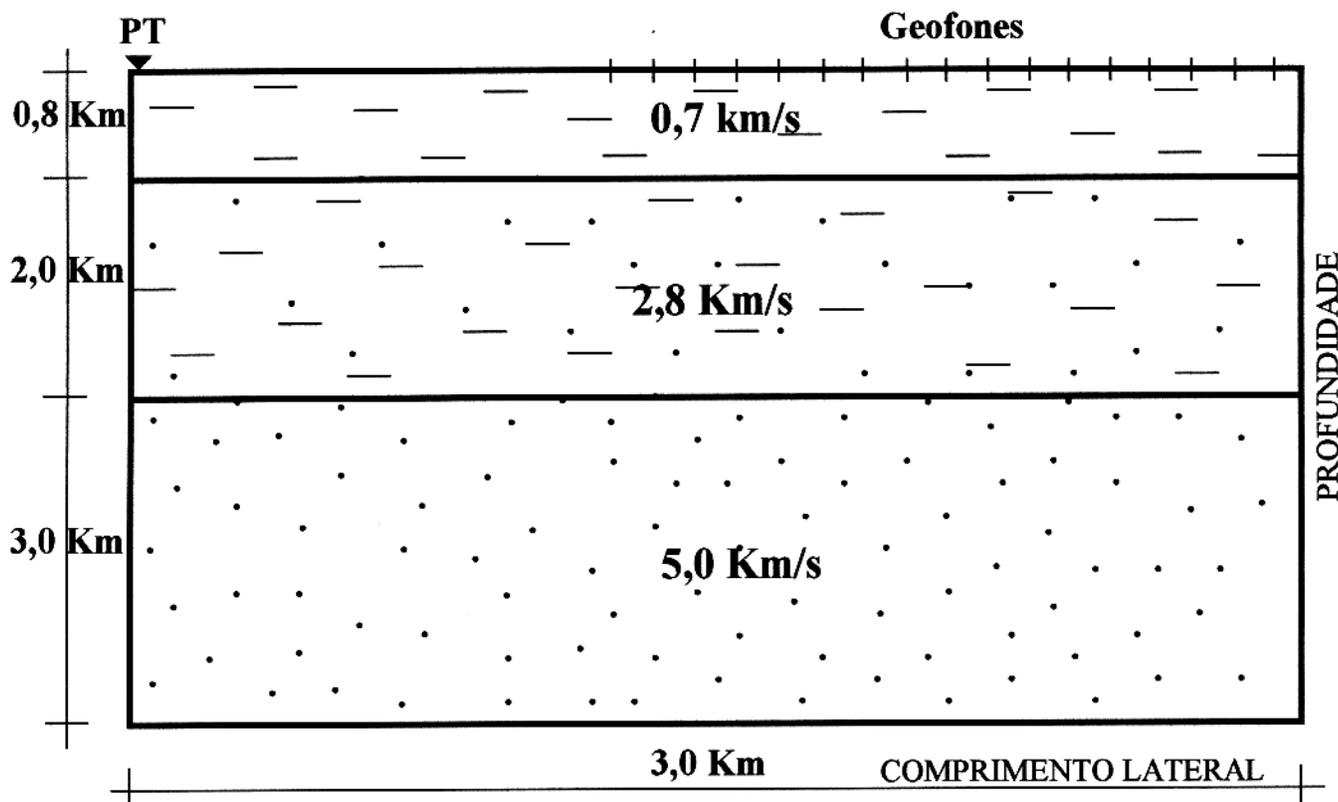


Figura 5.12: Modelo geofísico utilizado para a geração de um sismograma sintético com dois refletores.

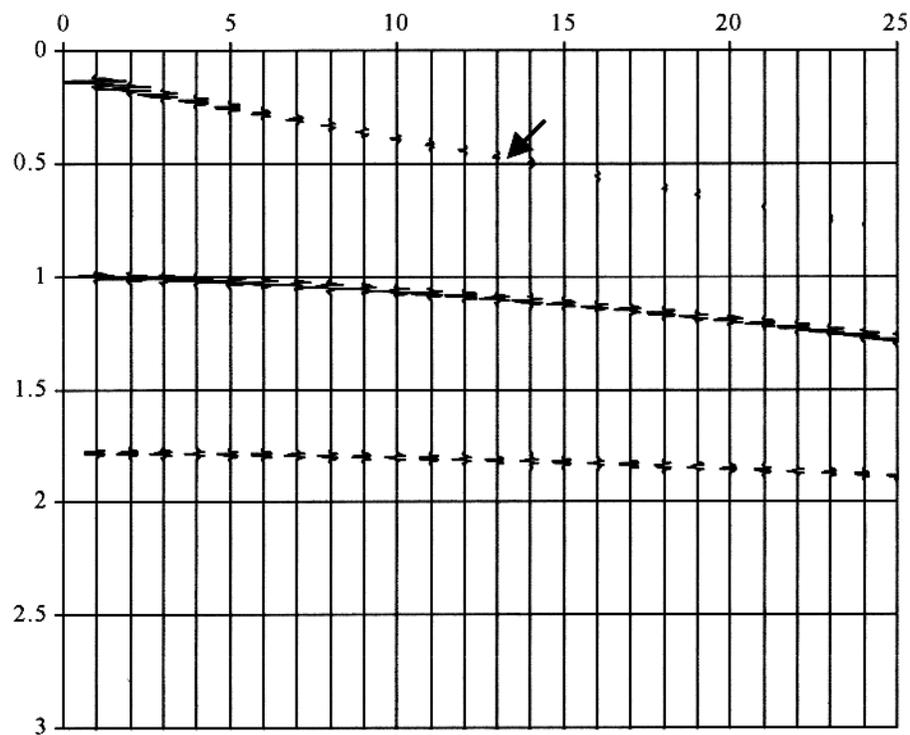


Figura 5.13: Sismograma sintético, com dois refletores, sem ruído.

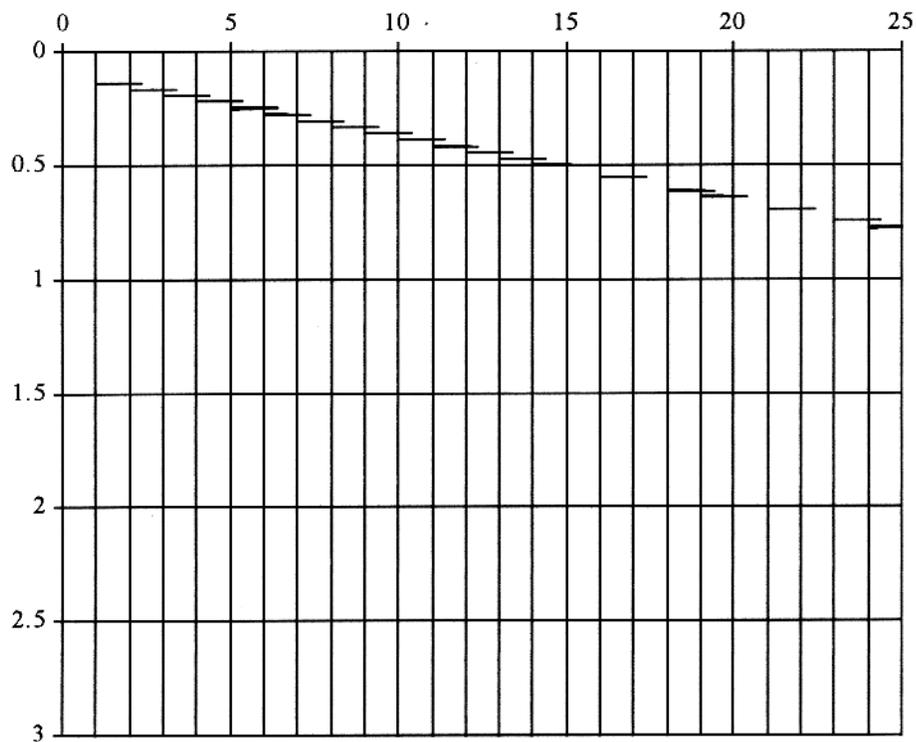


Figura 5.14: Primeiras quebras identificadas pela rede neural no sismograma sintético, com dois refletores, sem ruído.

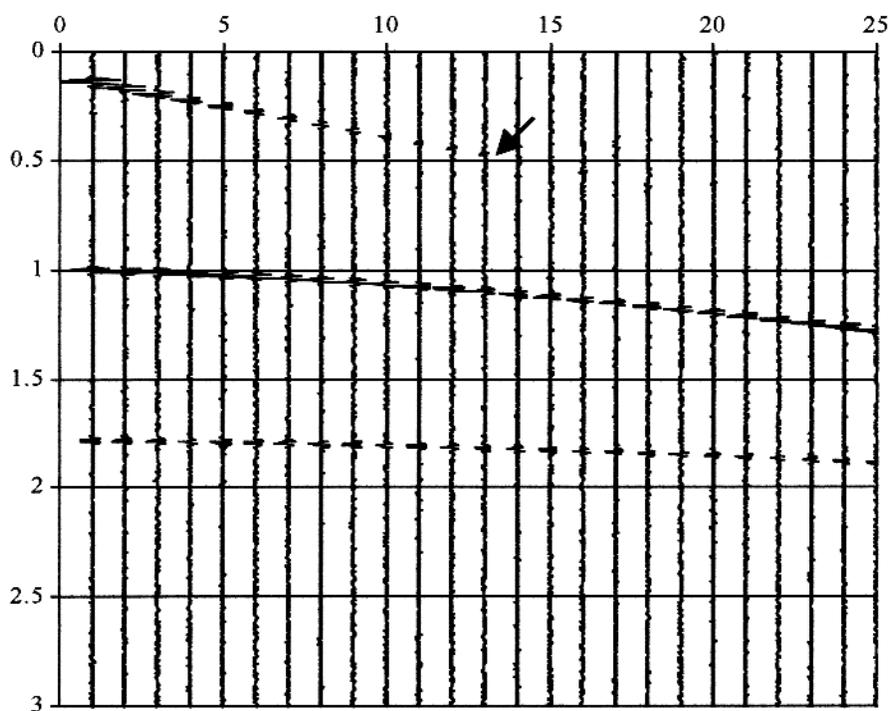


Figura 5.15: Sismograma sintético, com dois refletores, com ruído.

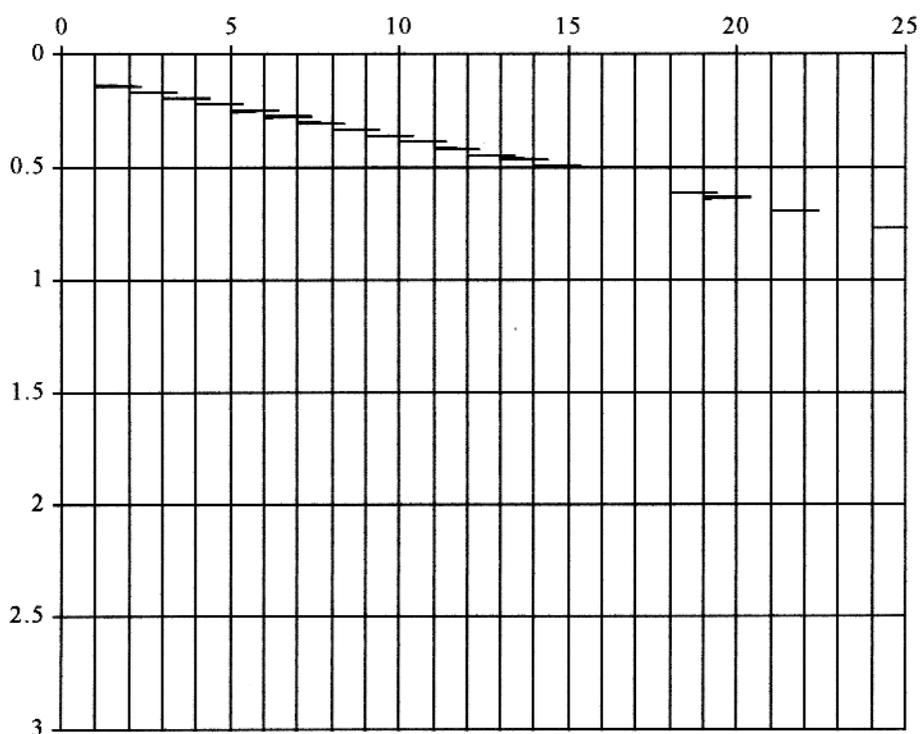


Figura 5.16: Primeiras quebras identificadas pela rede neural no sismograma sintético, com dois refletores, com ruído.

### 5.5.3- Sismograma com três refletores

Avaliaremos agora o comportamento da rede neural direta atuando sobre uma seção sísmica mais complexa, de modo a verificar a extensão do mapeamento realizado pela rede.

A seção que passaremos a analisar é um registro de tiro do tipo split-spread ( a fonte é posicionada no centro do arranjo de geofones), sobre um modelo geofísico com três refletores, conforme visto na Figura 5.17.

Vários conjuntos de treinamento foram testados na fase de aprendizagem da rede. Poderia-se supor que, num caso mais complexo, um conjunto de treinamento, como o adotado nos casos anteriores, não fosse capaz de ter incorporado informações suficientes para que a rede neural rastresse as primeiras quebras em todos os traços sísmicos. Fato este que não ocorreu. Segundo os nossos testes, independente da complexidade do modelo, com respeito à posição da fonte, número de refletores e o decaimento das amplitudes das primeiras quebras, a rede neural direta conseguiu efetivar a identificação das primeiras quebras a partir de um conjunto de treinamento bastante simples, como o dos casos anteriores.

O resultado da aplicação da rede neural direta sobre o sismograma da Figura 5.17 é visto na Figura 5.18, onde verifica-se a correta identificação dos tempos de chegada das primeiras quebras em todos os traços sísmicos.

Uma avaliação do comportamento da rede na identificação das primeiras quebras para o caso de uma seção mais complexa e com ruído, é feita a partir da seção sísmica mostrada na Figura 5.19, onde o ruído, considerado como uma série branca, tem amplitude máxima igual à um quinto da máxima amplitude registrada na seção. O resultado da aplicação da rede neural direta é visto na Figura 5.20, que mostra a identificação dos tempos de chegada das primeiras quebras, em cada traço, representados por *spikes*.

Na Figura 5.21 mostramos quatro instantes do treinamento da rede neural direta. Na cor azul é mostrado o conjunto de estímulos, na cor verde o conjunto de respostas desejadas e na cor vermelha a resposta da rede a cada instante do treinamento.

Finalmente, podemos concluir que a efetiva identificação das primeiras quebras pela rede neural direta independe da complexidade da seção sísmica. A condição necessária e suficiente é que o conjunto de treinamento contenha a informação da forma de onda representativa da primeira quebra nos traços sísmicos. Uma vez que essa forma de onda seja convenientemente escolhida

peço intérprete para a composição do conjunto de estímulos, a rede neural direta será capaz de identificar todas as primeiras quebras presentes na seção sísmica.

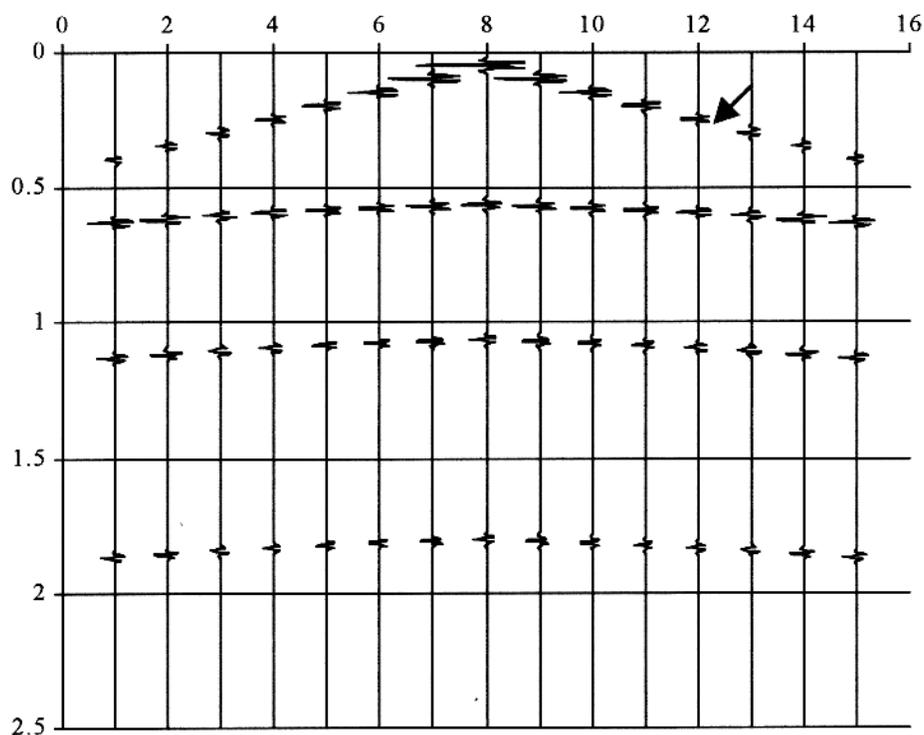


Figura 5.17: Sismograma sintético, com três refletores, sem ruído.

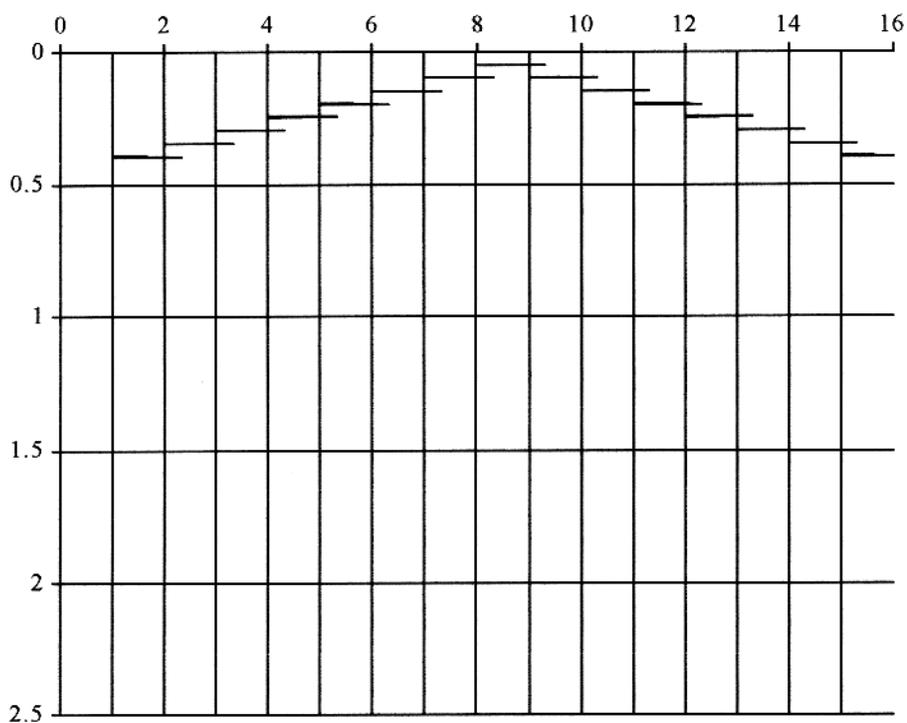


Figura 5.18: Primeiras quebras identificadas pela rede neural no sismograma sintético, com três refletores, sem ruído.

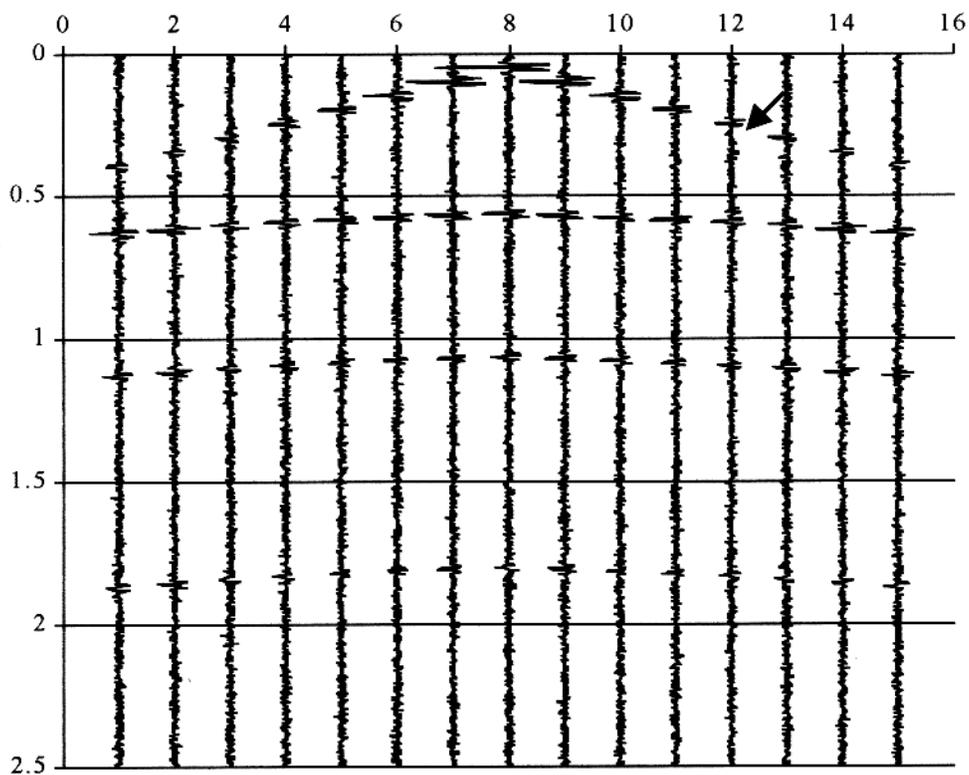


Figura 5.19: Sismograma sintético, com três refletores, com ruído.

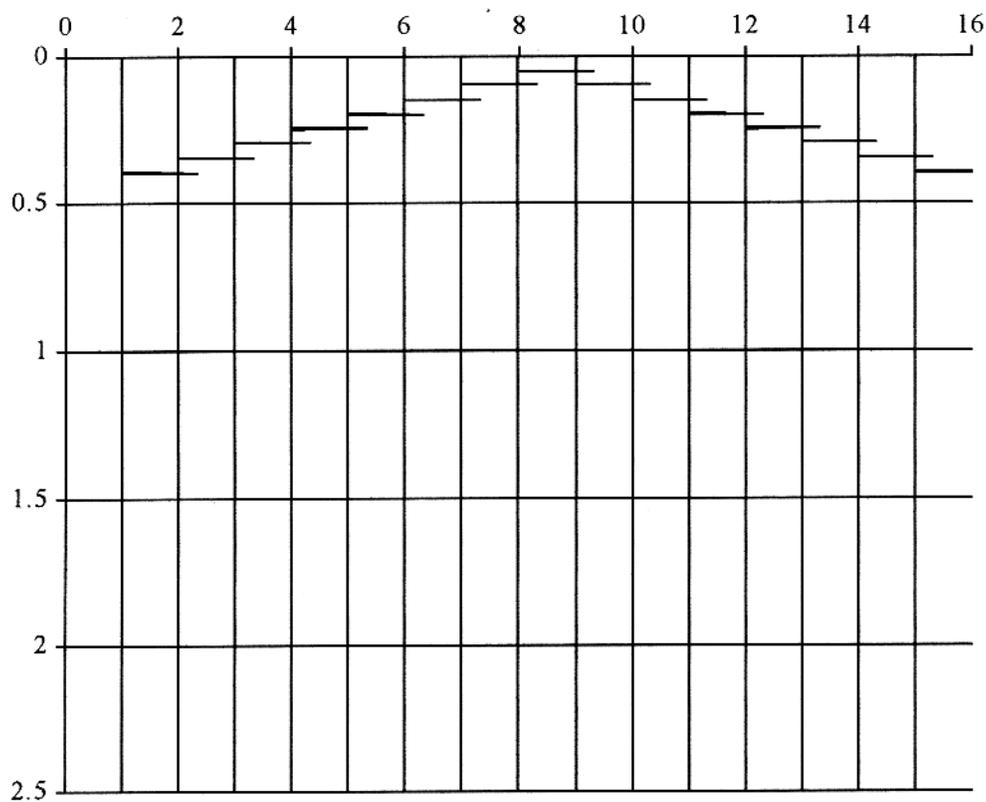


Figura 5.20: Primeiras quebras identificadas pela rede neural no sismograma sintético, com três refletores, com ruído.

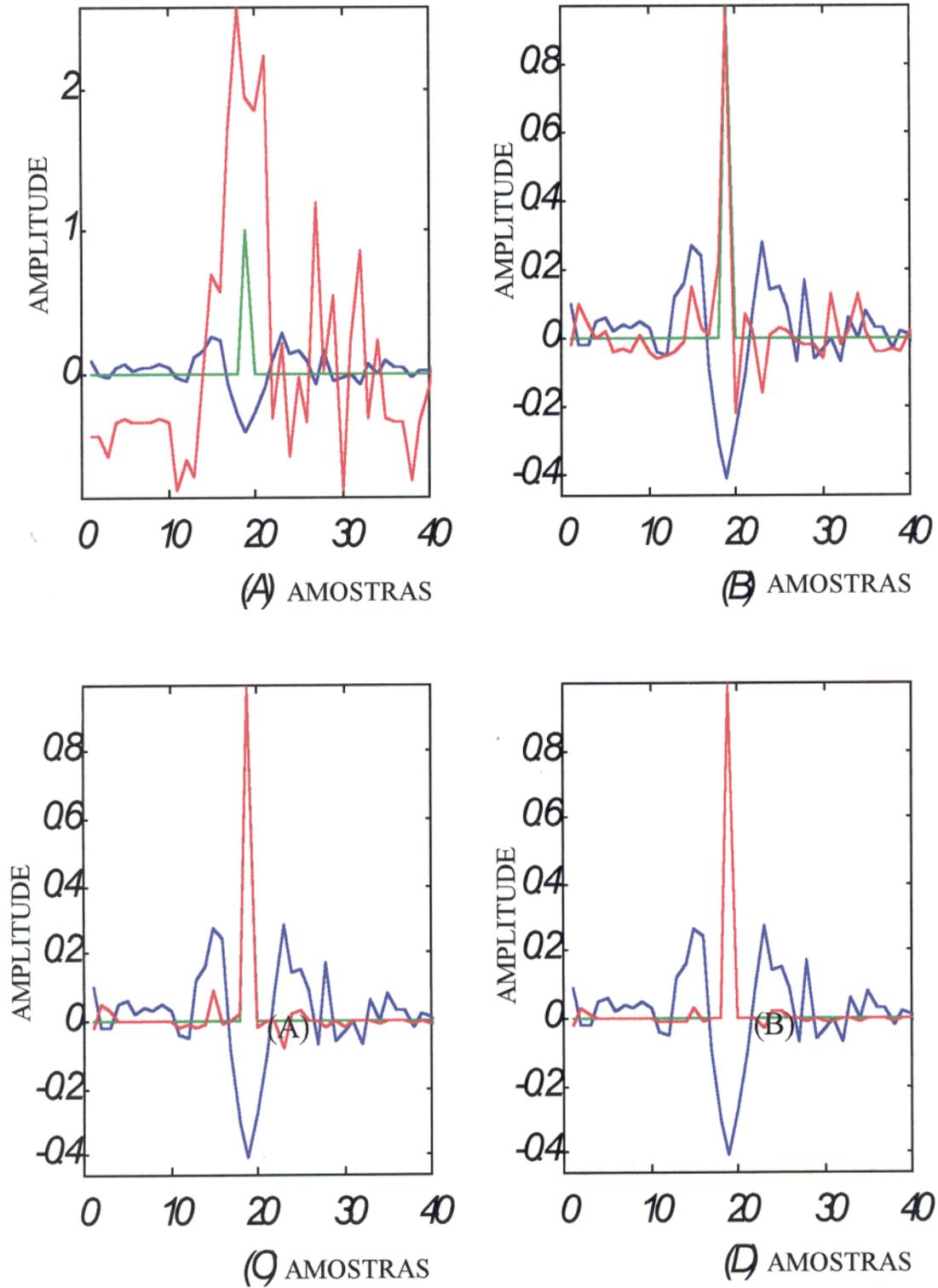


Figura 5.21: Exemplos de quatro instantes do treinamento da rede neural direta, mostrando, em azul, o conjunto de estímulos; em verde, o conjunto de respostas desejadas; e, em vermelho, a resposta da rede neural em cada instante do treinamento.

## 6- CONCLUSÕES

A simulação computacional do comportamento humano é uma tarefa extremamente difícil e porque não dizer em alguns casos impossível. A inventividade, a criatividade e a adaptabilidade do cérebro são características que os neurobiologistas e psiquiatras ainda buscam compreender em toda a sua extensão.

Para efeito desta dissertação não se cogitou, propriamente, da simulação do comportamento humano, mas sim da simulação do comportamento do sistema visual humano na forma de um problema de decisão bastante simples, baseado em regras bem definidas, que é a identificação das primeiras quebras no traço sísmico. No caso geral, um problema de decisão humano envolve uma série indeterminada de fatores psicológicos, culturais e emocionais, de natureza puramente qualitativa, que condicionam a escolha ou decisão e que, praticamente, inviabilizam uma versão computacional para este processo.

Uma das principais aplicações das redes neurais artificiais é a de simular, ainda que de um modo bastante simplificado, o comportamento do sistema visual humano em atividades como a do reconhecimento de padrões, classificações e processos de decisão. Com isto, a computação com uma rede neural artificial possibilita, de forma eficiente, a automação de alguns dos processos mentais que são peculiares ao ser humano, mas completamente estranhos aos algoritmos seqüenciais convencionais.

Entre os processos visuais realizados pelo intérprete sísmico, a identificação das primeiras quebras pode ser considerada como um dos mais simples. Assim, a simulação computacional do processo visual do intérprete pôde ser convenientemente tratada através de uma rede neural artificial multicamadas, na forma de um problema de decisão com, somente, duas alternativas possíveis. Após a inspeção de todo o traço sísmico, a decisão recai na identificação do evento representativo da primeira quebra e, conseqüentemente, na eliminação de todos os outros eventos anteriores e posteriores.

Assim, a arquitetura e o conjunto de treinamento da rede neural artificial, desenvolvidos para a identificação das primeiras quebras, mostraram-se eficazes para a satisfação de seus objetivos; apresentando algumas características que tornam este processamento passível de uma aplicação prática no auxílio ao intérprete sísmico:

1. Arquitetura extremamente otimizada, com grande eficiência computacional.
2. Possibilidade de utilização de conjuntos de treinamentos reduzidos e diferentes, implicando em uma maior eficiência na fase de treinamento.
3. Utilização do traço sísmico como única fonte de informação, prescindindo de qualquer tipo de pré-processamento.

A avaliação do comportamento ou do processamento de uma rede neural artificial, para a simulação de um processo visual de decisão, deve ser tratada de um modo completamente diferente daquele dado aos processos computacionais convencionais. A correção, ou a qualidade, do processamento com uma rede neural artificial não pode ser avaliada, simplesmente, em função da comparação da resposta fornecida pela rede neural com aquela que se considera como resposta correta.

O resultado da aplicação do processamento de uma rede neural deve, então, ser avaliado no sentido de obter-se uma verificação, mesmo que de forma qualitativa, da incorporação das informações contidas no conjunto de treinamento, estabelecido para a rede neural artificial, e a sua capacidade de utilização deste conhecimento em situações inéditas.

A utilização do conhecimento adquirido pela rede, durante a fase de treinamento, foi aqui avaliada pela razão entre a quantidade de informação fornecida através do conjunto de treinamento e a quantidade de novas informações obtidas por meio do seu processamento, ou seja, a informação fornecida sobre a forma ou formato de um único evento primeira quebra, retirado de um traço da seção sísmica, pode possibilitar, no processamento da rede, a identificação das primeiras quebras em todos os outros traços da seção sísmica.

A avaliação dos resultados da arquitetura de rede neural para a identificação das primeiras quebras deve também considerar a sua limitação, que é determinada pela acuidade visual do intérprete sísmico. Deste modo, a rede é perfeitamente capaz de identificar o evento primeira quebra, unicamente, se o intérprete também o for. Não implicando, no processamento da rede, a determinação das primeiras quebras, mas, simplesmente, a sua identificação.

As extensões imediatas dos assuntos abordados nesta dissertação seriam a avaliação da arquitetura de rede neural, aqui apresentada, para a edição de traços sísmicos e a determinação das velocidades intervalares, com dados sísmicos de poço (VSP).

## REFERÊNCIAS BIBLIOGRÁFICAS

- AMARI, S., 1977. " Neural theory of association and concept-formation." *Biological Cybernetics* 26, 175-185.
- AMARI, S., 1983. " Field theory of self-organizing neural nets." *IEEE Transactions on Systems, Man, and Cybernetics* SMC-13, 741-748.
- ANDERSON, J.A., Silverstein, J.W., Rite, S., Jones, R.S., 1977, Distinctive features, categorical perception and probability learning: Some applications of a neural model. *Psych. Ver.*, (84): 413-451.
- ANDERSON, R.G., and McMechan, G.A., 1989, Automatic editing of noisy seismic data: *Geophysical Prospecting*, 37, 875-892.
- ANDRADE, A.J.N., Fischetti, A.I., 1999, Automatic lithologies identification from well-logs. CD-ROM dos Anais do 6º Congresso Internacional da Sociedade Brasileira de Geofísica.
- ANDRADE, A.J.N., Luthi, S.M., 1997, Correlação estratigráfica de perfis de poço através de rede neuronal direta multicamadas. Anais do 5º Congresso Internacional da Sociedade Brasileira de Geofísica, volume I, 369-372.
- BARRON, A.R., 1992. " Neural net approximation." In *Proceedings of the Seventh Yale Workshop on Adaptive and Learning Systems*, pp. 69-72. New Haven, CT: Yale University.
- CARPENTER, G., Grossberg, S., 1986a , Adaptive resonance theory: Stable self-organization of neural recognition codes in response to arbitrary lists of input patterns. Eighth Annual Conference of the Cognitive Science Society, 45-62.
- CARPENTER, G., Grossberg, S., 1987a , Invariant pattern recognition and recall by an attentive self-organizing ART architecture in a nonstationary world. Proceedings of the *IEEE* First International Conference on Neural Networks: Vol. II, 737-746.
- CARPENTER, G., Grossberg, S., 1987b , A massively parallel architecture for a self-organizing neural pattern recognition machine. *Computer Vision, Graphics and Image Understanding*, 37, 54-115.
- CARPENTER, G., Grossberg, S., 1987c , Self-organization of stable category recognition codes for analog input patterns. Proceedings of the *IEEE* First International Conference on Neural Networks: Vol. II, 727-736.
- CARPENTER, G., Grossberg, S., 1987d , (Art.2): Self-organization of stable category recognition codes for analog input patterns, *Applied Optics*, 26, 4919-4930.

- CARPENTER, G.A., and S. Grossberg, 1987. "A massively parallel architecture for a self-organizing neural pattern recognition machine." *Computer Vision, Graphics, and Image Processing* **37**, 54-115.
- CARPENTER, G.A., M.A. Cohen, and S. Grossberg, 1987. Technical comments on "Computing with neural networks." *Science* **235**, 1226-1227.
- CERVENY, V., 1987. Ray Methods for Three-Dimensional Seismic Modelling, Petroleum Industry Course: Norwegian Institute for Technology.
- CERVENY, V. & Psencik, I., 1988. Program Seis88 – Numerical Modelling of Seismic Wave Fields in 2D Laterally Varying Layered Structures by the Ray Method. Czechoslovakia.
- CICHOCKI, A. & Unbehau, R., 1995. Neural Networks for Optimization and Signal Processing.
- COFFEEN, J.A., 1978. "Seismic Exploration fundamentals: the use of seismic techniques in finding oil." PennWell Publishing Company. Tulsa, Oklahoma.
- CYBENKO, G., 1988. *Approximation by Superpositions of a Sigmoidal Function*. Urbana: University of Illinois.
- CYBENKO, G., 1989. "Approximation by superpositions of a sigmoidal function." *Mathematics of Control, Signals, and Systems*, **2**, 303-314.
- EECKMAN, F.H. & Freeman, W.J., 1986. The Sigmoid Nonlinearity in Neural Computation: An Experimental Approach. Neural Networks for Computing.
- ERVIN, C.P., McGinnis, L.D., Otis, R.M., and Hall, M.L., 1983, Automated analysis of marine refraction data: a computer algorithm, *Geophysics*, **48**, 582-589.
- FISCHETTI, A.I., Andrade, A.J.N., 1999, Automatic first break picking in seismic traces by neural network. . CD-ROM dos Anais do 6º Congresso Internacional da Sociedade Brasileira de Geofisica.
- FUKUSSHIM, K., 1975." Cognitron: A self-organizing multilayered neural network." *Biological Cybernetics* **20**, 121-136.
- FUKUSHIMA, K., S. Miyake, and T. Ito, 1983. "Neocognitron: A neural network model for a mechanism of visual pattern recognition." *IEEE Transactions on Systems, Man, and Cybernetics* **SMC-13**, 826-834.
- GALLANT, A.R., and H. White, 1988. "There exist a neural network that does not make avoidable mistakes." *IEEE International Conference on neural Networks*, Vol.1, pp. 657-664, San Diego, CA.
- GELCHINSKY, B., and Shtivelman, V., 1983, Automatic picking of first arrivals and parameterization of traveltimes curves. *Geophysical Prospecting*, **31**, 915-928.

- GROSSBERG, S., 1964, The theory of embedding fields with applications to psychology and neurophysiology. Rockefeller Institute of Medical Research.
- GROSSBERG, S., 1970a, Neural pattern discrimination. *Journal of Theoretical Biology*, 27, 291-337.
- GROSSBERG, S., 1972b, Neural expectation: Cerebellar and retinal analogues of cells fired by unlearnable and learnable pattern classes. *Kybernetik*, 10, 49-57.
- GROSSBERG, S., 1972. " Neural expectation: Cerebellar and retinalanalogs of cells fires by learnable or unlearned pattern classes." *Kybernetic* 10, 49-57.
- GROSSBERG, S., 1973, Contour enhancement, short-term-memory, and constancies in reverberating network. *Studies in Applied Mathematics*, 52, 217-257.
- GROSSBERG, S., 1976b. " Adaptive pattern classification and universal recording: II. Feedback, expectation, olfaction, illusions." *Biological Cybernetics* 23, 187-202.
- GROSSBERG, S., 1978a. Competition, decision and consensus. *Journal of Analytic Applications*, 66, 470-493.
- GROSSBERG, S., 1987a, Competitive learning: from interactive activation to adaptive resonance. *Cognitive Science*, 11, 23-63.
- HAGEDOORN, 1959. " The plus-minus method of interpreting seismic refraction sections". *Geophysical Prospecting* , 7 (1959), pp. 158-182.
- HAYKIN, S., 1994. " Neural Networks - A Comprehensive Foundation." Macmillan College Publishing Company,. New York.
- HEBB, D.O., 1949. *The Organization of Behavior: A Neuropsychological Theory*. New york: Wiley.
- HECHT-NIELSEN, R., 1987. " Kolmogorov's mapping neural network existence theorem." 1<sup>st</sup> *IEEE International Conference on Neural Networks*, Vol. 3, pp. 11-14, San Diego, CA.
- HECHT-NIELSEN, R., 1988a . Applications of counterpropagation networks. *Neural Networks* (1): 131-140.
- HINTON, G.E., and S.J. Nowlan, 1987. " How learning can guide evolution." *Complex Systems* 1, 495-502.
- HORNIK, K., M. Stinchcombe, and H. White, 1989. " Multilayer feedforward network are universal approximators." *Neural Networks* 2, 359-366.

- HUSH, D. R., and B.G. Horne, 1993. " Progress in supervised neural networks: What's new since Lippman?" *IEEE Signal Processing Magazine* 10, 8-39.
- JACOBS, R. A., 1988, Increased Rates of Convergence through Learning Rate Adaptation. *Neural Networks*, 1.
- JANSSEN, P., P. Stoica, T. Söderström, and P. Eykhoff, 1988. " Model structure selection for multivariable systems by cross-validation." *International Journal of Control* 47, 1737-1758.
- KOHONEN, T., 1984, Self-organization and Associative Memory. Springer-Verlag, Berlin.
- KRAMER, A.H., and A. Sangivanni-Vincentelli, 1989. " Efficient parallel learning algorithms for neural networks." *In Advances in Neural Information Processing Systems 1* (D.S. Touretzky, ed.), pp. 40-48. San mateo, CA: Morgan Kaufmann.
- LECUN, Y., Boser, B., Denker, J.S., Henderson, D., Howard, R.E., Hubbard, W., and Jackel, L.D., 1990, Handwritten digit recognition with a back-propagation network, *Advances in Neural Information Processing Systems* 2, 396-404.
- LEE, Y., Oh, S., Kim, M., 1991, The effect of initial weights on premature saturation in back-propagation learning. *International Joint Conference on Neural Networks*. Vol. 1.
- McCLELLAND, J. et al., 1986. *The Appeal of Parallel Distributed Processing*
- McCORDOCK, P., 1979, *Machine who thinks*. W.H. Freeman and Company, p. 113.
- McCORMACK, M.D., Zaucha, D.E., and Dushek, D.W., 1993, First-break refraction event picking and seismic data trace editing using neural networks. *Geophysics*, 58, 67-78.
- McCULLOH, W.S., and W. Pitts, 1943. " A logical calculus of the ideas immanent in nervous activity." *Bulletin of Mathematical Biophysics* 5, 115-133.
- MINSKY, M.L., 1954. " Theory of neural-analog reinforcement systems and its application to the brain-model problem." Ph.D. Thesis, Princeton University, Princeton, NJ.
- MINSKY, M.L., and S.A. Papert, 1969. *Perceptrons*. Cambridge, MA: MIT Press.
- MURAT, M.E., and Rudman, A.J., 1992, Automated first arrival picking, A neural network approach, *Geophysical Prospecting*, 40, 587-604.
- NEFF, D.B., and Wyatt, S.B., 1986, Noise supression by the radial amplitude-slope rejection method. *Geophysics*, 51, 844-850.
- PARKER, D.B., 1985, Learning-logic: Casting the cortex of the human brain in silicon. Technical Report TR-44, Center for Computational Research in Economics and Management Science.

- PARKER, D.B., 1987, Optimal algorithms for adaptive networks: second order direct propagation and second order Hebbian learning. *IEEE*, 1<sup>st</sup> International Conference on Neural Networks, Vol. 2.
- PSENCIK, I., 1994, Introduction to seismic methods – Lecture Notes. PPPG/UFBa, Salvador-Ba.
- ROSENBLATT, F., 1958. " The Perceptron: A probabilistic model for information storage and organization in the brain." *Psychological Review* **65**, 386-408.
- ROSENBLATT, F., 1962. Principles of Neurodynamics. Spartan Books.
- RUMELHART, D.E., and J.L. McClelland, eds., 1986. *Parallel Distributed Processing: Explorations in the Microstructure of Cognition*, Vol. 1. Cambridge, MA: MIT Press.
- RUMELHART, D.E., Hinton, G.E., and Williams, R.J., 1986, Learning internal representation by error propagation. *Parallel Distributed Processing: Explorations in the microstructure of cognition*, Volume 1, 318-362.
- SACKINGER, E., Boser, B.E., Bromley, J., LeCun, Y., and Jackel, L.D., 1992, Application of the ANNA neural network chip to high-speed character recognition, *IEEE Transactions on Neural Networks* **3**, 498-505.
- SEJNOWSKI, T.J., Yuhas, B.P., Goldenstein, M.H., Jr., and Jenkins, R.E., 1990, Combining visual and acoustic speech signals with a neural network improves intelligibility. *Advances in Neural Information Processing Systems* **2**, 232-239.
- SHERIFF, Robert E.; Geldart, Lloyd P., 1982. "Exploration Seismology-second edition". Cambridge University Press.
- SHEPHERD, G.M., and C. Koch, 1990. " Introduction to synaptic circuits." In *The Synaptic Organization of the Brain* (G.M. Sheperd, ed.). pp. 3-31. New York: Oxford University press.
- SONTAG, E.D., 1992. " Feedback stabilization using two-hidden-layers nets." *IEEE Transactions on neural Networks* **3**, 981-990.
- SPAGNOLINI, U., 1991, Adaptive picking of refracted first arrivals. *Geophysical Prospecting*, **39**, 293-312.
- STEINBUCH, K., and Piske, U.A.W., 1963. " Learning matrices and their applications." *IEEE Transactions on Electronic Computers* **EC-12**, 846-862.
- STEINBUCH, K., Zendeh, F., 1963, Self-connecting translator circuits. *Information Processing*, 359-365.
- VON NEWMANN, J., 1958. *The Computer and the Brain*. Yale University Press.

- WERBOS, P.J., 1974. " Beyond regression: New tools for prediction and analysis in the behavioral sciences." Ph.D. Thesis, Harvard University, Cambridge, MA.
- WIDROW, B., and M.E. Hoff, Jr., 1960. " Adaptive switching circuits." *IRE WESCON Convention Record*, pp. 96-104.
- WIDROW, B., Groner, G., Hu, M., Smith, F., Specht, D., Talbert, L., 1963, Pratical applications for adaptive data-processing system. WESCON Convention Record, Section 11.4.
- WIDROW, B., Smith, F., 1964, Pattern-recognizing control system. *Computer and Information Sciences*, 288-317.
- WIDROW, B. & Hoff, M.E., 1966. Adaptive swiching circuits. *IRE WESCOM Convention Record*.
- WIDROW, B., Mantey,P., Griffiths, L., 1967, Adaptive antenna systems. *Proceedings of the IEEE*, 55, 2143-2159.
- WIDROW, B., and S.D. Stearns, 1985. *Adaptive Signal Processing*. Englewood Cliffs, NJ: Prentice-Hall.
- WIDROW, B., N.K. Gupta, and S. Maitra, 1973. " Punish/reward: Learning with a critic in adaptive treshold systems." *IEEE Transactions on Systems, Man, and Cybernetics SMC-3*, 455-465.
- WIDROW, B., J.R. Glover, Jr., J.M. McCool, J. Kaunitz, C.S. Williams, R.H. Hearn , J.R. Zeidler, J. Dong, Jr., and R.C. Goodlin, 1975. " Adaptive noise cancelling: Principles and applications." *Proceedings of the IEEE* 63, 1692-1716.
- WIDROW, B., and S.D. Stearns, 1985. " Adaptive Signal Processing." *Englewood Cliffs, NJ: Prentice-Hall*.
- WIDROW, B., 1988, The original adaptive neural network broom balancer. *International Symposium on Circuits and Systems*, 351-357.