

UNIVERSIDADE FEDERAL DE UBERLÂNDIA
CENTRO DE CIÊNCIAS EXATAS E TECNOLOGIA
CURSO DE PÓS-GRADUAÇÃO EM ENGENHARIA MECÂNICA

**TÉCNICAS DE CARACTERIZAÇÃO DE EXCITAÇÕES EM MÁQUINAS
ROTATIVAS**

Tese apresentada à Universidade Federal de Uberlândia por:

ADEMYR GONÇALVES DE OLIVEIRA

em 25/06/1999 como parte dos requisitos necessários para a obtenção
do título de Doutor em Engenharia Mecânica

Banca examinadora:

Prof. Dra. Kátia Lucchesi Cavalca Dedini - UNICAMP
Prof. Dr. Robson Pederiva - UNICAMP
Prof. Dr. Francisco Paulo Lépore Neto - UFU- (Orientador)
Prof. Dr. Carlos Roberto Ribeiro - UFU
Prof. Dr. Marcus Antônio Viana Duarte - UFU

Aos meus pais, Graciema e Thomaz.

Agradecimentos

Ao Professor Francisco Paulo Lépore Neto, pelo apoio durante o período de orientação do trabalho.

Ao Professor Valder Steffen Jr., coordenador do Projeto FAPEMIG TEC 544/91, pelo apoio financeiro e colaboração científico.

À CAPES - Fundação Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - e à FAPEMIG - Fundação de Apoio à Pesquisa do Estado de Minas Gerais - pelo apoio financeiro.

Índice

Lista de Figuras e Tabelas	IV
Resumo	VI
Abstract	VII
Capítulo 1 - Introdução	9
1.1 - Definição do problema	9
Capítulo 2 - Fundamentos de Dinâmica dos Rotores	13
2.1 - Introdução	13
2.2 - Elemento Disco	14
2.3 - Elemento árvore.....	16
2.3.1- Energia Cinética.....	17
2.3.2 - Energia potencial.	21
2.4 - Mancais.....	26
2.5 - Excitação por desbalanceamento	27
2.6 - Equações de Movimento do Rotor	30
2.7 - Resposta de Sistema Rotativo à Excitações: Método Pseudo Modal	30
2.7.1 - Solução da equação homogênea	32
2.7.2 - Solução Particular: Resposta à Excitações	32
Capítulo 3 - Introdução à Análise de Sinal	35
3.1 - Introdução	35
3.2 - Revisão de Análise de Sinal	36
3.3 - Análise no Domínio do Tempo	37
3.4 - Análise no domínio da frequência	39
3.4.1 - Função densidade de potência	39
3.4.2 - Transformada Rápida de Fourier (FFT).....	41
3.4.3 - Transformada de Fourier de dados discretos.....	43
3.4.4 - Correlação e Autocorrelação usando FFT	44
3.4.5 - Densidade Espectral de Potência (PSD) estimada pela FFT	45
3.4.6 - Filtragem dos Dados	47
3.5 - Transformada de Wavelet.....	49
3.5.1 - Módulo e Fase da Transformada de Wavelet	51
3.5.2 - Aplicação: Estimação da frequência Natural e Razão de Amortecimento.	52
Capítulo 4 - Fundamentos de Redes Neurais	55
4.1- Introdução	55
4.2- O que são e quais as aplicações das redes neurais?	57
4.3- Conceitos básicos	58

4.3.1- Unidade de processamento: neurônio	58
4.3.2- Ligações entre neurônios: sinapses.	58
4.3.3- Função de ativação	59
4.4- Fundamentos	61
4.4.1- Topologia da Rede Neural	62
Rede neural unidirecional (feed-forward networks)	62
Rede neural recorrente (feedback networks)	62
Rede Neural Multicamadas Unidirecional	62
4.4.2- Treinamento das Redes Neurais	63
A Regra Delta	63
O Algoritmo Back-propagation	65
A Regra Delta Generalizada	65
4.5- Trabalhando com Back-propagation	67
4.5.1- Ajuste dos pesos aplicado à função de ativação sigmoial	68
4.5.2- Taxa de aprendizado e o termo de inércia	69
4.5.3 - Deficiências do método de Back-propagation.	69
Paralisia neural	70
Mínimo local	70
Aprendizagem viciada.	70
Ajuste inadequado	71
4.6 - Algoritmos avançados	73
4.7 - Redes Neurais Auto-organizadas	74
4.7.1 - Classificação de padrões	75
A aprendizagem competitiva	75
Quantização vetorial	77
4.8 - Redes Recorrentes	78
4.8.1 - A Rede de Hopfield	79
4.9 - Redução Dimensional: Método Estatístico	80
4.9.1 - Componentes Principais (cp)	81
4.9.2 - Escalonamento e alterações numéricas	83
4.9.3 - Componentes principais de um grupo de centróides	84
4.9.4 - Interpretação do método de compactação de dados.	86
Capítulo 5 - Bancada de Dinâmica de Rotação	89
5.1 - Introdução	89
5.2 - Bancada de Dinâmica de Rotação	89
5.2.1 - Modelo computacional	92
5.2.2 - Caracterização experimental do rotor	93
5.2.3 - Caracterização experimental da estrutura	95

5.2.4 - Comparações de resultados	96
5.3- Sistema de Medição	100
Capítulo 6 - Fontes de Excitação	105
6.1 - Excitação por desbalanceamento.	106
6.2 - Excitação por força de natureza magnética.	106
6.3 - Excitação dos mancais por força assíncrona.	108
Capítulo 7 - Teste da Metodologia Proposta	111
7.1 - Introdução.	111
7.2 - Modelo de 3 Graus de Liberdade	111
7.3 - Localização da fonte de excitação	113
7.4 - Sensibilidade da Rede Neural à Presença de Ruído.	113
7.5 - Influência da Compactação dos dados no Funcionamento da Rede Neural.	114
Capítulo 8 - Resultados e Conclusões	117
8.1 - Rede neural de identificação de excitação por desbalanceamento.	117
8.2 - Rede neural de identificação de excitação assíncrona no mancais.	118
8.3 - Rede neural de identificação de excitações por forças magnéticas.	120
8.4 - Rede neural de classificação.	122
8.5 -Conclusões.	123
Referências Bibliográficas	131
ANEXO I - Componentes da Bancada	137
ANEXO II - Modelo de Elementos Finitos do Rotor - ANSYS	153
ANEXO III - Modelo de Elementos Finitos da Estrutura da Bancada - ANSYS	157
ANEXO IV - Sistema de 1 GDL: Caracterização do Amortecedor.	165
ANEXO V - Função de transferência do rotor.	169
ANEXO VI - Função de transferência da estrutura da bancada.	173
ANEXO VII - Curvas de calibração do sistema de medição.	179
ANEXO VIII - Programa MKDX.for e CUTGDL.for	185
ANEXO IX - Programa de Rede Neural (BACKPMA.C)	191
ANEXO X - Programa COMPACT.M	203

Lista de Figuras e Tabelas

Figura (1.1) - Procedimento de análise dos dados.	11
Figura (2.1) - Elemento de disco: Sistema de Eixos Coordenados	15
Figura (2.2) - Elemento de árvore	16
Figura (2.3) - Deslocamentos nodais nos planos XY e ZY.	18
Figura (2.4) - Secção transversal de elemento árvore	21
Figura (2.5) - Configuração de um mancal	26
Figura (2.6) - Excitação por desbalanceamento	28
Figura (2.7) - Massa de desbalanceamento.	28
Figura (3.1) - Densidade de potência teórica e densidade de potência mensurável.	41
Figura (3.2) - Fenômeno aliasing	44
Figura (3.3) - Funções janela mais usadas.	48
Figura (3.4) - Leakage das funções janela.	49
Figura (4.1) - Modelo gráfico de um neurônio	58
Figura (4.2) - Modelo gráfico de uma sinapse dupla	59
Figura (4.3) - Funções de ativação.	59
Figura (4.4) - Rede unidirecional multicamadas.	63
Figura (4.5) - O ajuste dos pesos.	69
Figura (4.6) - Arquitetura apropriada. 5x10x1.	72
Figura (4.7) - Arquitetura 5x5x1: fenômeno do subajuste.	72
Figura (4.8) - Arquitetura 5x40x1: fenômeno do sobreajuste.	72
Figura (4.9) - Redes auto-organizadas	75
Figura (4.10) - Classes e centróides	76
Figura (4.11) - Subespaços disjuntos	77
Figura (4.12) - Rede neural anelar	79
Figura (4.13) - Dois componentes principais	81
Figura (4.14) - Influência do número de componentes principais	83
Figura (4.15) - Formação dos componentes principais.	85
Figura (5.1) - Rotor: elemento móvel da bancada.	90
Figura (5.2) - Sistema de acionamento	90
Figura (5.3) - Suspensão superior do rotor.	91
Figura (5.5) - Vista da configuração da suspensão inferior do rotor.	94
Figura (5.4) - Vista geral da bancada de teste de rotores	92
Figura (5.6) - Diagrama de Campbell do rotor, obtido experimentalmente.	95
Figura (5.7) - Ensaio das frequências naturais da estrutura	96
Figura (5.8) - Modos de vibrar do rotor.	97

Figura (5.9) - Primeiros modos de flexão da estrutura.	98
Figura (5.10) - Travessas de suporte dos sensores de proximidade.	99
Figura (5.11) - Função de transferência da estrutura, excitação: A1 e medição: A2. ...	100
Figura (5.12) - Diagrama de blocos do sistema de medição.	101
Figura (5.13) - Sistema de aquisição de dados	102
Figura (5.14) - Diagrama de blocos do programa de aquisição de dados.	102
Figura (5.15) - Interface gráfica do programa de aquisição.	103
Figura (5.16) - Circuito amplificador inversor.	104
Figura (6.1) - Posicionamento dos sensores.	105
Figura (6.2) - Posicionamento das massas de desbalanceamento	106
Figura (6.3) - Detalhe da excitação de natureza magnética.	107
Figura (6.4) - Força magnética aplicada ao disco 2, amplitude 9 N, freqüência 10 Hz.	107
Figura (6.5) - Detalhe da excitação assíncrona nos mancais.	108
Figura (7.1) - Fluxograma do procedimento de caracterização das excitações	111
Figura (7.2) - Sistema mecânico de 3 graus de liberdade (gdl).	112
Figura (7.3) - Função resposta em freqüência: sistema 3 GDL.	112
Figura (8.1) - Fluxograma geral do processo automatizado de diagnostico de defeitos.	126
Tabela (5.1) - Valores de rigidezes dos mancais.	93
Tabela (5.2) - Valores do coeficiente de amortecimento dos mancais.	94
Tabela (5.3) - Freqüências naturais do rotor para $\Omega = 0$	96
Tabela (5.4) - Freqüências naturais da estrutura.	99
Tabela (7.1) - Localização do ponto da aplicação força: na presença de ruídos.	113
Tabela (7.2) - Localização do ponto da aplicação força: dados compactados.	114
Tabela (8.1) - Qualificação das redes neurais: desbalanceamento.	118
Tabela (8.2) - Qualificação das redes neurais: excitação assíncrona	119
Tabela (8.3) - Qualificação das redes neurais: excitação magnética.	121
Tabela (8.4) - Rede neural de classificação.	122
Tabela (8.5) - Rede neural de classificação: dados não compactados.	123

Resumo

Este trabalho apresenta uma contribuição ao estudo e caracterização de forças de excitações aplicadas em maquina rotativas, utilizando redes neurais a partir do sinal de vibração medido no equipamento. A metodologia desenvolvida é utilizada para classificar a excitação entre quatro estados de funcionamento da maquina: operação normal, excitação por força de desbalanceamento, excitação por força assíncrona e excitação por força magnética. Uma vez encontrado o tipo de excitação é possível determinar o seu ponto de aplicação e suas características de amplitude e frequência.

Aplicando uma técnica estatística de compactação dos sinais medidos, foi possível treinar com sucesso redes neurais com número de neurônios menores e conseqüentemente, com menor custo computacional.

A eficiência e robustez das arquiteturas propostas, de redes neurais, foram avaliadas para diferentes níveis de compactação de dados e de ruídos aditivos utilizando simulação numérica de um modelo de sistemas vibratório de três graus de liberdade.

A metodologia foi validada numa bancada experimental que representa uma máquina rotativa de rotor flexível.

Redes neurais. Dinâmica de rotação. Forças de excitação.

Summary

This work presents a contribution to the study and characterization of excitements forces applied in rotative machine, using neural networks starting from the measured vibration sign in the equipment. The developed methodology is used to classify the excitement among four states of operation conditions: normal operation, unbalance excitement force, asynchronous excitement force and magnetic excitement force. For the each excitement class it is possible to determine its application point and its amplitude and frequency characteristics.

Applying a technique statistics data compress of the measured signs, it was possible to training with success neural networks with smaller number of neurons, consequently with smaller computational cost.

The efficiency and robustness of the architectures proposals, of neural networks, they were appraised for different levels of data compress and of addictive noises using numeric simulation of a vibratory model of systems of three degrees of freedom.

The methodology was validated in a experimental apparatus that represents rotative machine whit a flexible rotor.

Neural networks. Rotor dynamics. Excitement forces.

Capítulo 1

Introdução

Modernamente existe uma grande diversidade de aplicações de máquinas rotativas envolvendo tanto equipamentos de grande porte, como as unidades de geração de energia elétrica a partir de fontes hidráulicas ou térmicas, e que operam em velocidades relativamente baixas, como também equipamentos de pequenas dimensões, como por exemplo bombas criogênicas, ultra-centrífugas e turbo compressores, que giram em altas velocidades.

O problema de desbalanceamentos e outros, que provocam níveis excessivos de vibração e até instabilidade, tem exigido o desenvolvimento de técnicas de controle destas vibrações e técnicas de monitoração de seu desempenho dinâmico, para garantir o seu funcionamento dentro de padrões de segurança. Há de se considerar ainda o alto valor econômico destes equipamentos e a necessidade de mantê-los em operação com alto nível de confiabilidade.

O monitoramento, identificação e diagnóstico de desgaste e falhas, em componentes críticos de equipamentos, bem como a evolução dos defeitos operacionais, exigem o desenvolvimento e a automatização de metodologias precisas, robustas, de resposta rápida, com baixo custo operacional.

Conhecendo os sinais de vibração de uma dada máquina, antes e após o aparecimento de um defeito (fonte perturbadora), como determinar a localização e a característica desse defeito? Qual a dificuldade de automatizar este diagnóstico? Estas são a chave da questão.

O grupo de Dinâmica da UFU, dentro do escopo do projeto TEC 544/91 financiado pela FAPEMIG, desenvolveu um instrumental capaz de permitir o estudo dos fenômenos relacionados com a dinâmica de máquinas rotativas, caracterizado pelo projeto e construção de uma bancada de ensaios de máquinas rotativas, a qual permite uma ampla variedade de experimentos.

O objetivo desse trabalho é consolidar e desenvolver técnicas computacionais e experimentais capazes de proceder a análise do comportamento dinâmico de máquinas rotativas, com ênfase na caracterização de excitações, com vistas à monitoração de sua operação e à sua manutenção preditiva.

Neste contexto, as redes neurais artificiais são apresentadas como uma poderosa ferramenta que reúne características desejáveis como:

- a abstração do modelo matemático do sistema,
- a capacidade de mapear sistemas não-lineares,
- a capacidade de inferir resultados,
- a insensibilidade a ruídos e
- a possibilidade de adaptar-se às mudanças na planta via procedimentos de auto-aprendizagem.

Além dessas características, as redes neurais têm custo relativamente baixo de implementação e processamento rápido, atributos que sugerem um alto grau de automatização do processo de manutenção preditiva. Isso é o que diz a teoria. A eficácia desses atributos serão então testados experimentalmente em dinâmica de rotação.

Esse trabalho tem por objetivo, propor uma metodologia para automatizar o processo de monitoração e diagnóstico de defeitos em máquinas rotativas, envolvendo as etapas de aquisição de sinais de vibração, pré-processamento e compressão dos dados via redes neurais, com vistas a identificação de excitações quanto a sua localização e caracterização.

É proposta uma técnica de qualificação de rede neural com base na média do erro quadrático.

A metodologia desenvolvida é validada por testes de simulação de um sistema vibratório linear de 3 graus de liberdade e por resultados provenientes de ensaios experimentais, realizados numa bancada constituída por um rotor flexível, com diferentes tipos de excitações aplicadas em pontos estratégicos, bem como por técnicas de análise de sinais já consagradas.

A figura (1.1) apresenta o procedimento a ser proposto para a caracterização de excitações numa máquinas rotativas.

Os sinais de vibração medidos no equipamento, representados no domínio da frequência, contem as informações sobre as excitações presentes na máquina. Tais sinais são pré-processados e geram os dados de entrada para uma rede neural de classificação, treinada para reconhecer 4 estados diferentes do equipamento: funcionamento normal, desbalanceamento do rotor, excitações nos mancais e excitações eletromagnéticas.

Uma vez determinado o tipo de excitação, os dados são enviados a uma rede neural especialista que caracteriza os parâmetros da força excitadora e da sua posição na máquina.

Em linhas gerais este trabalho constitui-se das seguintes etapas:

- 1 - Projetar, construir e montar uma bancada de teste de rotores.
- 2 - Instrumentar e construir dispositivos necessários para este fim.
- 3 - Identificar os parâmetros dinâmicas da bancada, tais como frequências e modos de vibração, bem como os parâmetros característicos dos mancais, com vistas a seleção dos pontos apropriados para excitação e medição, identificação da faixa de frequências mais apropriada para excitação e operação da bancada.
- 4 - Implementação do modelo computacional da bancada, com o objetivo de simular o seu funcionamento permitindo, juntamente com os dados experimentais, gerar banco de dados voltados para o treinamento das redes neurais.
- 5 - Definir um procedimento experimental capaz de reproduzir excitações do tipo harmônicas, periódicas e magnéticas.

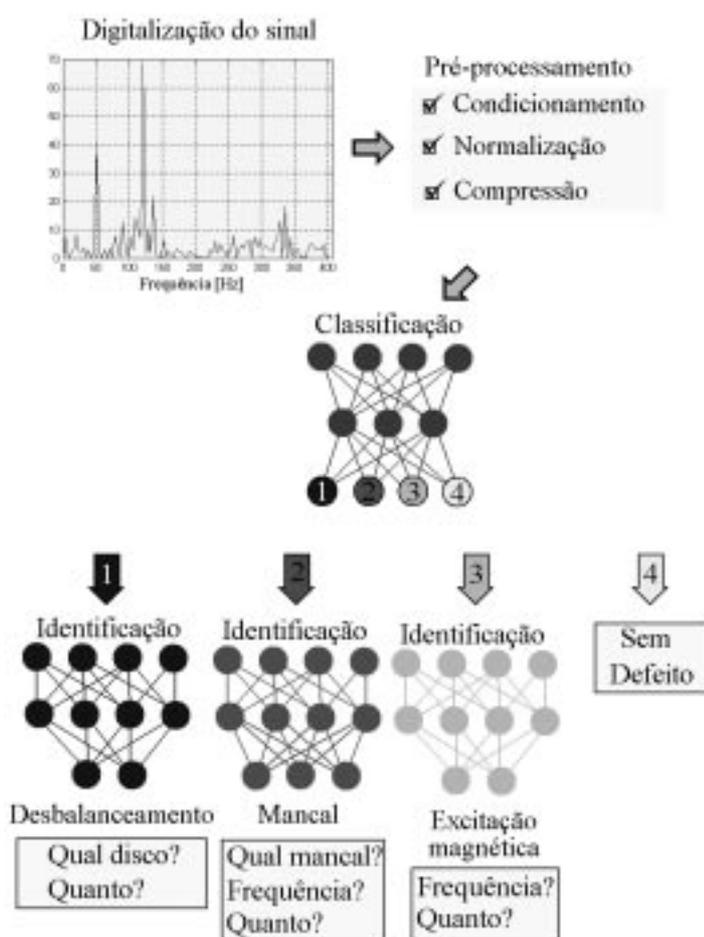


Figura (1.1) - Procedimento de análise dos dados.

- 6 - Adquirir e condicionar os sinais através de processos tais como normalização, compactação, ou mudança de unidade.
- 7 - Definir uma arquitetura da rede neural apropriada, voltada à caracterização e a localização da fonte excitadora.
- 8 - Aplicar excitações conhecidas em pontos estratégicos da bancada e, a partir de seus sinais, tentar identifica-las através da metodologia desenvolvida.

No capítulo 2 é apresentado o equacionamento de rotores através do método dos elementos finitos, onde é modelado os elementos de árvore, disco rígido e mancais, a partir das expressões da energia potencial, da energia cinética e do trabalho virtual de cada um dos elementos. Esta formulação é a base necessária para entender o funcionamento interno dos programas de elementos finitos disponíveis no mercado.

No capítulo 3 é feita uma revisão das principais ferramentas de análise de sinais, nos domínios do tempo e da frequência. É feita uma revisão da ferramenta matemática que é a transformada de wavelet.

No capítulo 4 é feita uma revisão sobre redes neurais, com ênfase no algoritmo de treinamento supervisionado, o back propagation. É mostrado ainda uma metodologia baseada na variância, que possibilita a condensação de um grande volume de dados.

No capítulo 5 é discutido o projeto, construção, montagem e caracterização de uma bancada, destinada ao estudo de problemas relacionados à dinâmica de rotores. É mostrada a operacionalização de um sistema de aquisição e processamento de dados, condicionadores de sinais, excitadores, dispositivos de fixação de sensores e de dispositivos de segurança.

No capítulo 6 são apresentados algumas técnicas de geração de excitações por forças de desbalanceamento, forças assíncronas nos mancais e forças de natureza magnéticas, em máquinas rotativas. São abordados aspectos quanto instrumentação, a segurança do operador e a integridade dos instrumentos e do rotor.

No capítulo 7 é testada a metodologia proposta utilizando uma implementação de um sistema vibratório de 3 graus de liberdade. Uma rede neural é treinada para identificar a localização da posição de uma força excitadora. É feito um estudo a respeito da robustez das redes neurais quanto à presença de ruído nos dados de treinamento.

No capítulo 8, é apresentado os resultados de uma rede neural treinada para classificar o tipo de excitação presente na máquina rotativa. Após a identificação do tipo de defeito presente na máquina, outra rede neural especialista caracteriza a força excitadora quanto à sua amplitude, frequência e posição ao longo do rotor. Várias redes neurais são testadas para avaliar a influência da arquitetura da rede neural, a qualidade dos dados quanto à localização dos sensores e taxa de compactação.

Capítulo 2

Fundamentos de Dinâmica dos Rotores

2.1 - Introdução

A resposta dinâmica de estruturas mecânicas tem sido alvo de investigação de um grande número de pesquisadores. Dentre os objetivos de tal esforço, pode-se destacar a caracterização das frequências, modos próprios de vibrar e o estudo da resposta a diferentes excitações.

A compreensão de tais características permite ao pesquisador alterar conscientemente, determinados parâmetros de um dado projeto, tais como dimensões e materiais, de modo a obter ou modificar características específicas do seu comportamento, vistas à redução dos custos de fabricação, aumento da confiabilidade, durabilidade e performance.

A formulação de um modelo matemático que representa de maneira razoável, um sistema real não é tarefa simples requer o conhecimento prévio de parâmetros de projeto como dimensões e dados de materiais. É necessário ainda dispor de um elenco de hipóteses simplificadoras que viabilizam o modelo numérico sem contudo descaracterizar seu comportamento. Por essa razão é quase sempre necessário a validação do modelo teórico com o modelo experimental, seja em escala real ou reduzida.

Define-se graus de liberdade como sendo o número de variáveis independentes de deslocamentos e rotações, necessários para a representação do modelo analítico do sistema mecânico.

O modelo analítico é constituído de um conjunto de equações diferenciais que descrevem o movimento do sistema nas direções de seus graus de liberdade, e são obtidas usando as Leis de Newton, relações de tensão deformação e equações de Lagrange (Craig, Jr. e Roy, R., 1981).

A literatura apresenta farto material sobre a obtenção das equações de movimento de rotores, destacando-se o método dos elementos finitos e o método das matrizes de transferência (Steffen Jr., V. 1981) (Lallement, Lecoanet e Steffen; 1982) (Lalanne, Queau e Tran, 1982) (Ruhl, R. e Booker, J. 1972) (Berthier, Ferraris e Lalanne, 1983).

O método dos elementos finitos, trata um sistema contínuo discretizando-o num conjunto de elementos que, individualmente, são considerados contínuos.

Na sua essência, o método dos elementos finitos expressa o deslocamento de qualquer ponto do sistema contínuo em termos dos deslocamentos de um conjunto finito de pontos, ditos pontos nodais, multiplicados por uma dada função de interpolação. Este método produz resultados

satisfatórios no estudo de problemas estruturais, sendo utilizado em um grande número de programas comerciais voltados para a análise estática e dinâmica de sistemas mecânicos.

No mercado de softwares para engenharia existem vários programas destinados a análise dinâmica de sistemas mecânicos. Dentre os mais conhecidos destacam-se o ANSYS™, NASTRAN™, ROTOR e o ADINA™. Estes são códigos computacionais validados e quando operados por profissionais capacitados, produzem resultados confiáveis, tornando-se ferramentas que permitem um grande número de análises a respeito do comportamento do modelo.

Será mostrado aqui a técnica de elementos finitos aplicadas à modelagem matemática de rotores flexíveis (Minh, T. D., 1981) (Lalanne, M., et al, 1987) (Engels, R. C., 1991) e o método modal de resolução do sistema de equações diferenciais (Bathe, K. e Wilson, E. L., 1986).

O modelo físico de uma máquina rotativa pode ser concebido pela combinação de elementos classificados como discos rígidos, árvores flexíveis e mancais.

As equações do movimento para os elementos de disco, árvores e mancais são obtidas aplicando as equações de Lagrange, equação (2.1), nas expressões das energias potencial e cinética dos elementos de disco, árvore e mancal.

$$\frac{d}{dt} \left(\frac{\partial T}{\partial \dot{q}_i} \right) - \frac{\partial T}{\partial q_i} + \frac{\partial U}{\partial q_i} = Q_i \quad (2.1)$$

Onde q_i é a i -ésima coordenadas generalizada, Q_i é a i -ésimo esforço generalizado, T e U são energias cinética e potencial respectivamente.

2.2 - Elemento Disco

O elemento de disco é assumido como sendo rígido e é caracterizado pela sua energia cinética.

A figura (2.1) ilustra o elemento de disco com o sistema de eixos coordenados.

R_0 e R_1 são respectivamente o referencial inercial e o referencial móvel, no disco rígido, com origem localizada no centro geométrico do disco.

$R_1(x, y, z)$ e $R_0(X, Y, Z)$ são relacionados através das variáveis Ψ , θ e ϕ , ângulos em torno dos eixos z , x e y respectivamente. A equação (2.2) mostra a matriz de transformação entre os sistemas R_0 e R_1 .

$$\begin{Bmatrix} x \\ y \\ z \end{Bmatrix} = \begin{bmatrix} \cos \Psi \cos \phi - \sin \Psi \sin \theta \sin \phi & \sin \Psi \cos \phi + \cos \Psi \sin \theta \sin \phi & -\cos \theta \sin \theta \\ -\sin \Psi \cos \theta & \cos \Psi \cos \theta & \sin \theta \\ \cos \Psi \sin \phi + \sin \Psi \sin \theta \cos \phi & \sin \Psi \sin \phi - \cos \Psi \sin \theta \cos \phi & \cos \theta \cos \phi \end{bmatrix} \begin{Bmatrix} X \\ Y \\ Z \end{Bmatrix} \quad (2.2)$$

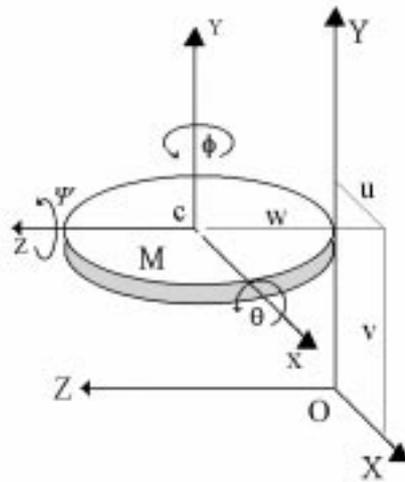


Figura (2.1) - Elemento de disco: Sistema de Eixos Coordenados

Como já foi dito, o disco é assumido como sendo constituído por um elemento rígido e suas matrizes de massa e a giroscópica são obtidas a partir da expressão da energia cinética, que é dada pela equação (2.3), onde w é o vetor de velocidade instantânea de rotação do referencial móvel (R_1) dada pela equação (2.4).

$$T_D = \frac{1}{2} M_D (\dot{u}^2 + \dot{w}^2) + \frac{1}{2} I_{Dx} (\dot{\theta}^2 + \dot{\Psi}^2) + \frac{1}{2} I_{Dy} (\Omega^2 + 2\Omega\dot{\Psi}\theta) \quad (2.3)$$

$$w = \begin{Bmatrix} \dot{w}_x \\ \dot{w}_y \\ \dot{w}_z \end{Bmatrix} = \begin{Bmatrix} \dot{\theta} \cos \phi - \dot{\Psi} \cos \theta \sin \phi \\ \dot{\phi} + \dot{\Psi} \sin \theta \\ \dot{\theta} \sin \phi + \dot{\Psi} \cos \theta \cos \phi \end{Bmatrix} \quad (2.4)$$

O movimento do elemento disco é descrito por duas translações u e w , e duas rotações θ e Ψ . Logo, define-se o vetor de coordenadas generalizadas como mostra a equação (2.5).

$$q = \begin{Bmatrix} u \\ w \\ \theta \\ \Psi \end{Bmatrix} \quad (2.5)$$

Aplicando a equação de Lagrange, equação (2.1), nas equações da energia cinética, equação (2.3), e vetor de coordenadas generalizadas, equação (2.5), obtém-se a matriz massa M_D e a matriz giroscópica G_D do modelo.

$$M_D = \begin{vmatrix} M_D & 0 & 0 & 0 \\ 0 & M_D & 0 & 0 \\ 0 & 0 & I_{DX} & 0 \\ 0 & 0 & 0 & I_{DY} \end{vmatrix} \quad (2.6)$$

$$G_D = \begin{vmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -I_{DY} \\ 0 & 0 & I_{DY} & 0 \end{vmatrix} \quad (2.7)$$

2.3 - Elemento árvore.

O elemento árvore é caracterizado pela energia cinética e energia elástica de deformação, considerando que ele possui secção transversal circular e um comprimento L, como mostra a figura (2.2).

A secção transversal é admitida ser linearmente variável. São considerados os efeitos da inércia de rotação da secção transversal (Rayleigh) e o efeito de cisalhamento da secção transversal (Timoshenko).

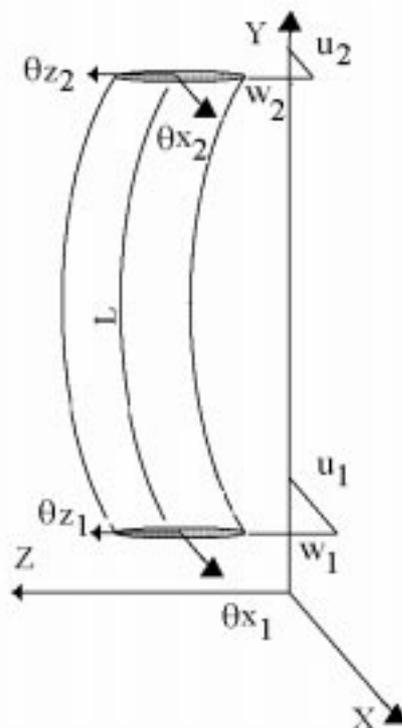


Figura (2.2) - Elemento de árvore

2.3.1- Energia Cinética

A formulação geral da energia cinética T_A , para o elemento árvore, é uma extensão do conceito de energia cinética T_D , para o elemento disco, calculado para um certo comprimento L. Integrando a equação (2.3) obtém-se:

$$T_A = T_1 + T_2 + T_3 + T_4 \quad (2.8)$$

onde:

$$T_1 = \frac{1}{2} \rho S \int_0^L (\dot{u}^2 + \dot{w}^2) dy \quad \text{energia cinética de viga submetida à flexão}$$

$$T_2 = \frac{1}{2} \rho I \int_0^L (\dot{\Psi}^2 + \dot{\theta}^2) dy \quad \text{efeito de Rayleigh}$$

$$T_3 = \frac{1}{2} \rho I_p L \Omega^2 \quad \text{energia cinética de rotação } (\Omega \text{ constante})$$

$$T_4 = \rho I_p \Omega \int_0^L \dot{\Psi} \theta dy \quad \text{efeito giroscópico}$$

$$I_p \quad \text{inércia polar da seção transversal}$$

$$I \quad \text{inércia da seção transversal}$$

$$\rho \quad \text{densidade volumétrica}$$

$$S \quad \text{é a área da seção transversal}$$

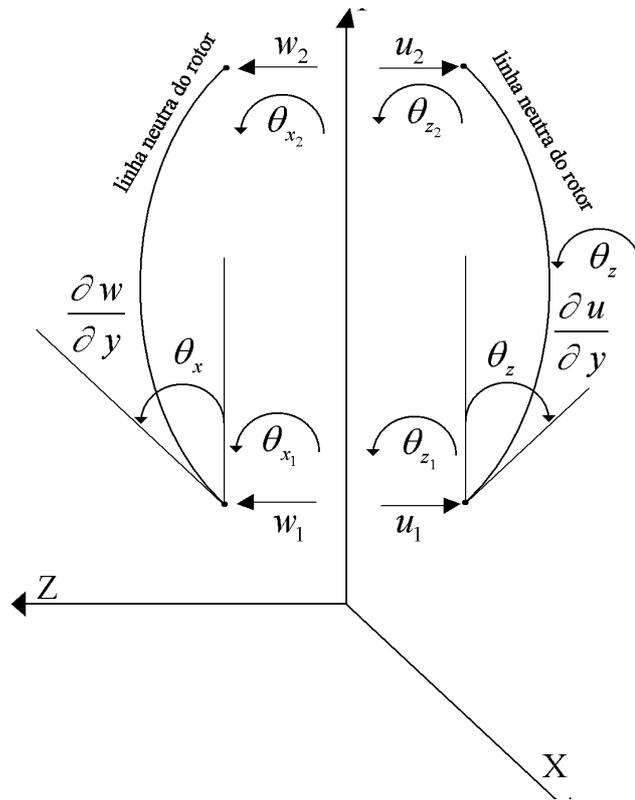


Figura (2.3) - Deslocamentos nodais nos planos XY e ZY.

A figura (2.3) mostra a projeção dos deslocamentos nodais nos planos XY e ZY, de uma árvore orientada na direção do eixo Y e compreendida entre dois nós. Cada nó possui dois graus de liberdade, totalizando quatro graus de liberdade para o elemento de árvore. A hipótese assumida é a de que o deslocamento nodal pode ser descrito por uma equação polinomial de terceira ordem, tal como segue:

$$w = c_1 + c_2 y + c_3 y^2 + c_4 y^3 \quad (2.9)$$

Derivando a equação (2.9) com relação ao deslocamento y tem-se:

$$\theta_x = \frac{-\partial w}{\partial y} = c_2 + 2 c_3 y + 3 c_4 y^2 \quad (2.10)$$

Resolvendo a equação (2.9) e equação (2.10) para as condições de contorno w_1 e θ_1 em $y=0$ e w_2 e θ_2 em $y=L$ ficam determinadas as constantes c_1 , c_2 , c_3 e c_4 , resultando as equações do deslocamento e rotação ao longo do comprimento da árvore.

$$w(y) = \left[1 - \frac{3y^2}{L^2} + \frac{2y^3}{L^3} \mid y - \frac{2y^2}{L} + \frac{y^3}{L^2} \mid \frac{3y^2}{L^2} - \frac{2y^3}{L^3} \mid \frac{y^3}{L^2} - \frac{y^2}{L} \right] \begin{bmatrix} w_1 \\ \theta_1 \\ w_2 \\ \theta_2 \end{bmatrix} \quad (2.11)$$

A equação 2.11 pode ser reescrita como:

$$w(y) = N_2 \delta w \quad (2.12)$$

Logo:

$$\Psi(y) = \frac{\partial N_2(y)}{\partial y} \delta w \quad (2.13)$$

O mesmo procedimento pode ser empregado para a obtenção de $u(y)$ e $w(y)$ (deslocamento e rotação no plano ZY), observando a convenção dos sentidos mostrados no plano ZY da figura (2.3).

$$u(y) = \left[1 - \frac{3y^2}{L^2} + \frac{2y^3}{L^3} \quad \middle| \quad -y + \frac{2y^2}{L} - \frac{y^3}{L^2} \quad \middle| \quad \frac{3y^2}{L^2} - \frac{2y^3}{L^3} \quad \middle| \quad \frac{y^2}{L} - \frac{y^3}{L^2} \right] \begin{Bmatrix} x \\ y \\ z \end{Bmatrix} = \begin{Bmatrix} u_1 \\ \Psi_1 \\ u_2 \\ \Psi_2 \end{Bmatrix} \quad (2.14)$$

Ou

$$u(y) = N_1(y) \delta u \quad (2.15)$$

Logo:

$$\Psi(y) = \frac{\partial N_1(y)}{\partial y} \delta u \quad (2.16)$$

Inserindo as expressões das variáveis u , w , θ e Ψ na equação da energia cinética do elemento de árvore, equação (2.8), e integrando ao longo do seu comprimento, de 0 até L, as equações de Lagrange para coordenadas generalizadas ficam assim definidas.

$$\frac{d}{dt} \left(\frac{\partial T_A}{\partial \dot{q}_i} \right) - \frac{\partial T_A}{\partial q_i} = [M_A + M_s] \{\ddot{q}\} + [\Omega G_A] \{\dot{q}\} \quad (2.17)$$

onde M_A e M_s representam respectivamente a matriz de massa e inércia e G_A é a matriz giroscópica do elemento árvore, sendo representadas pelas equações 2.18, 2.19 e 2.20 respectivamente, onde I é a inércia da secção transversal da árvore, L é o comprimento do elemento, ρ é a densidade do material, S é a área da secção transversal e I_p é a inércia polar da secção transversal do elemento árvore.

$$M_A = \frac{\rho S L}{420} \begin{array}{|c|c|c|c|c|c|c|c|} \hline 156 & 0 & 0 & -22L & 54 & 0 & 0 & 13L \\ \hline 0 & 156 & 22L & 0 & 0 & 54 & -13L & 0 \\ \hline 0 & 22L & 4L^2 & 0 & 0 & 13L & -3L^2 & 0 \\ \hline -22L & 0 & 0 & 4L^2 & -13L & 0 & 0 & -3L^2 \\ \hline 54 & 0 & 0 & -13L & 156 & 0 & 0 & 22L \\ \hline 0 & 54 & 13L & 0 & 0 & 156 & -22L & 0 \\ \hline 0 & -13L & -3L^2 & 0 & 0 & -22L & 4L^2 & 0 \\ \hline 13L & 0 & 0 & -3L^2 & 22L & 0 & 0 & 4L^2 \\ \hline \end{array} \quad (2.18)$$

$$M_S = \frac{\rho I}{30L} \begin{array}{|c|c|c|c|c|c|c|c|} \hline 36 & 0 & 0 & -3L & -36 & 0 & 0 & -3L \\ \hline 0 & 36 & 3L & 0 & 0 & -36 & 3L & 0 \\ \hline 0 & 3L & 4L^2 & 0 & 0 & -3L & -L^2 & 0 \\ \hline -3L & 0 & 0 & 4L^2 & 3L & 0 & 0 & -L^2 \\ \hline -36 & 0 & 0 & 3L & 36 & 0 & 0 & 3L \\ \hline 0 & -36 & -3L & 0 & 0 & 36 & -3L & 0 \\ \hline 0 & 3L & -L^2 & 0 & 0 & -3L & 4L^2 & 0 \\ \hline -3L & 0 & 0 & -L^2 & -3L & 0 & 0 & 4L^2 \\ \hline \end{array} \quad (2.19)$$

$$G_A = \frac{\rho I_p}{30L} \begin{array}{|c|c|c|c|c|c|c|c|} \hline 0 & -36 & -3L & 0 & 0 & 36 & -3L & 0 \\ \hline 36 & 0 & 0 & -3L & -36 & 0 & 0 & -3L \\ \hline 3L & 0 & 0 & -4L^2 & -3L & 0 & 0 & L^2 \\ \hline 0 & 3L & 4L^2 & 0 & 0 & -3L & -L^2 & 0 \\ \hline 0 & 36 & 3L & 0 & 0 & -36 & 3L & 0 \\ \hline -36 & 0 & 0 & 3L & 36 & 0 & 0 & 3L \\ \hline 3L & 0 & 0 & L^2 & -3L & 0 & 0 & -4L^2 \\ \hline 0 & 3L & -L^2 & 0 & 0 & -3L & 4L^2 & 0 \\ \hline \end{array} \quad (2.20)$$

2.3.2 - Energia potencial.

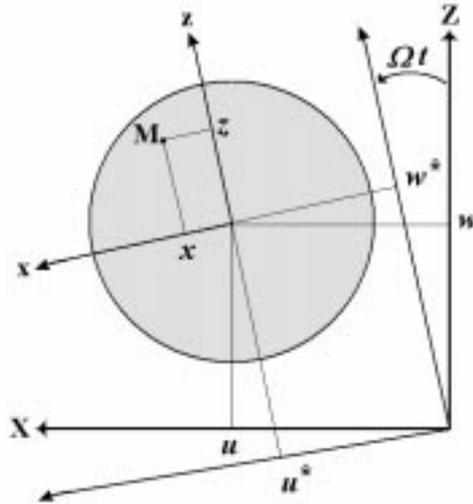


Figura (2.4) - Secção transversal de elemento árvore

Da resistência dos materiais tem-se a expressão da deformação \mathcal{E} de um ponto $M(x, y)$ contido na secção transversal de uma viga submetida a flexão, como mostra a figura (2.4).

$$\mathcal{E} = \underbrace{-x \frac{\partial^2 u^*}{\partial y^2} - z \frac{\partial^2 w^*}{\partial y^2}}_{\varepsilon_l} + \underbrace{\frac{1}{2} \left(\frac{\partial u^*}{\partial y} \right)^2 + \frac{1}{2} \left(\frac{\partial w^*}{\partial y} \right)^2}_{\varepsilon_{nl}} \quad (2.21)$$

Onde ε_l e ε_{nl} são respectivamente a parcela linear e não linear da expressão da deformação. Da figura (2.4), obtêm-se as seguintes relações::

$$\begin{aligned} u^* &= u \cos \Omega t - w \sin \Omega t \\ w^* &= u \sin \Omega t + w \cos \Omega t \end{aligned} \quad (2.22)$$

A relação entre tensão σ e a deformação é dada pela equação (2.23) onde E denota o modulo de Young.

$$\sigma = E\varepsilon \quad (2.23)$$

A expressão clássica da energia de deformação de viga submetida a carregamento radial é dada pela equação (2.24).

$$U_1 = \frac{1}{2} \int_{\tau} \varepsilon' \sigma d\tau \quad (2.24)$$

Conhecendo as equações (2.21) e (2.23), a equação (2.24) pode ser particularizada para representar a energia de deformação do elemento árvore, como mostrada na equação (2.25).

$$U_1 = \frac{E}{2} \int_{\tau} \varepsilon_l^2 d\tau + \int_{\tau} E \varepsilon_l \varepsilon_{nl} d\tau + \frac{E}{2} \int_{\tau} \varepsilon_{nl}^2 d\tau \quad (2.25)$$

O segundo termo da equação anterior é nulo para o elemento de árvore, devido à simetria com relação aos eixos coordenados x e z .

Desconsiderando os termos de segunda ordem, a equação (2.24) torna-se

$$U_1 = \frac{E}{2} \int_0^L \int_s \left(-x \frac{\partial^2 u^*}{\partial y^2} - z \frac{\partial^2 w^*}{\partial y^2} \right)^2 ds dy \quad (2.26)$$

Da resistência dos materiais define-se os momentos de inércia de área:

$$\begin{aligned} I_x &= \int_s z^2 ds \\ I_z &= \int_s x^2 ds \end{aligned} \quad (2.27)$$

Expandindo a equação (2.26) tem-se:

$$U_1 = \frac{E}{2} \int_0^L \int_s \left(x^2 \left(\frac{\partial^2 u^*}{\partial y^2} \right)^2 + z^2 \left(\frac{\partial^2 w^*}{\partial y^2} \right)^2 + 2xz \frac{\partial^2 u^*}{\partial y^2} \frac{\partial^2 w^*}{\partial y^2} \right) ds dy \quad (2.28)$$

Considerando a condição de simetria da secção transversal do elemento árvore, implica na nulidade do terceiro termo da equação (2.28). Lembrando-se ainda das equações (2.27), a equação (2.28) pode ser reescrita como:

$$U_1 = \frac{E}{2} \int_0^L I_z \left(\frac{\partial^2 u^*}{\partial y^2} \right)^2 dy + \frac{E}{2} \int_0^L I_x \left(\frac{\partial^2 w^*}{\partial y^2} \right)^2 dy \quad (2.29)$$

Uma nova parcela de energia de deformação U_2 deve ser considerada quando existe um carregamento axial constante F_0 .

$$U_2 = \frac{F_0}{2} \int_0^L \varepsilon_l dy + \frac{F_0}{2} \int_0^L \varepsilon_{nl} dy \quad (2.30)$$

Lembrando-se da simetria da secção transversal no plano xz anteriormente assumida, implica na nulidade do primeiro termo da equação (2.30).

Da equação (2.21) tem-se a expressão de ε_{nl} , logo a equação (2.29) torna-se:

$$U_2 = \frac{F_0}{2} \int_0^L \left(\frac{\partial u^*}{\partial y} \right)^2 dy + \frac{F_0}{2} \int_0^L \left(\frac{\partial w^*}{\partial y} \right)^2 dy \quad (2.31)$$

A energia de deformação total do elemento de árvore é dada por $U_A = U_1 + U_2$, logo:

$$U_A = \frac{E}{2} \int_0^L I_z \left(\frac{\partial^2 u^*}{\partial y^2} \right)^2 dy + \frac{E}{2} \int_0^L I_x \left(\frac{\partial^2 w^*}{\partial y^2} \right)^2 dy + \frac{F_0}{2} \int_0^L \left(\frac{\partial u^*}{\partial y} \right)^2 dy + \frac{F_0}{2} \int_0^L \left(\frac{\partial w^*}{\partial y} \right)^2 dy \quad (2.32)$$

Da condição de simetria da secção transversal do elemento árvore tem-se que $I_x = I_z$. Da equação (2.22) tem-se as expressões de u^* e w^* . A equação (2.32) pode então ser reescrita como:

$$U_A = \frac{EI}{2} \int_0^L \left(\left(\frac{\partial^2 u}{\partial y^2} \right)^2 + \left(\frac{\partial^2 w}{\partial y^2} \right)^2 \right) dy + \frac{F_0}{2} \int_0^L \left(\left(\frac{\partial u}{\partial y} \right)^2 + \left(\frac{\partial w}{\partial y} \right)^2 \right) dy \quad (2.33)$$

As equações (2.11) e (2.14) definem os deslocamentos nodais w e u , logo tem-se que:

$$U_A = \frac{1}{2} \delta u^t K_1 \delta u + \frac{1}{2} \delta w^t K_2 \delta w + \frac{1}{2} \delta u^t K_3 \delta u + \frac{1}{2} \delta w^t K_4 \delta w \quad (2.34)$$

Onde:

$$K_1 = EI \int_0^t \frac{\partial^2 N_1(y)}{\partial y^2} \frac{\partial^2 N_1(y)}{\partial y^2} dy \quad K_2 = EI \int_0^t \frac{\partial^2 N_2(y)}{\partial y^2} \frac{\partial^2 N_2(y)}{\partial y^2} dy \quad (2.35)$$

$$K_3 = F_0 \int_0^t \frac{\partial N_1(y)}{\partial y} \frac{\partial N_1(y)}{\partial y} dy \quad K_4 = F_0 \int_0^t \frac{\partial N_2(y)}{\partial y} \frac{\partial N_2(y)}{\partial y} dy$$

K_1 e K_2 são as matrizes clássicas de rigidez à flexão.

K_3 e K_4 são matrizes de rigidez devido ao carregamento axial ditas matrizes rigidez das condições iniciais.

Resolvendo as equações (2.35), resultam:

$$K_1 = \frac{EI}{L^3} \begin{vmatrix} 12 & -6L & -12 & -6L \\ -6L & 4L^2 & 6L & 2L^2 \\ -12 & 6L & 12 & 6L \\ -6L & 2L^2 & 6L & 4L^2 \end{vmatrix} \quad (2.36)$$

$$K_2 = \frac{EI}{L^3} \begin{vmatrix} 12 & 6L & -12 & 6L \\ 6L & 4L^2 & -6L & 2L^2 \\ -12 & -6L & 12 & -6L \\ 6L & 2L^2 & -6L & 4L^2 \end{vmatrix} \quad (2.37)$$

$$K_3 = \frac{F}{30L} \begin{vmatrix} 36 & -3L & -36 & -3L \\ -3L & 4L^2 & 3L & -L^2 \\ -36 & 3L & 36 & 3L \\ -3L & -L^2 & 3L & 4L^2 \end{vmatrix} \quad (2.38)$$

$$K_4 = \frac{F}{30L} \begin{vmatrix} 36 & 3L & -36 & 3L \\ 3L & 4L^2 & -3L & -L^2 \\ -36 & -3L & 36 & -3L \\ 3L & -L^2 & -3L & 4L^2 \end{vmatrix} \quad (2.39)$$

Considerando-se ainda o elemento de árvore modelado como viga de Timoshenko, ou seja, computando o efeito do cisalhamento na secção transversal da viga, provocado pela sua flexão, uma energia potencial adicional U^* pode ser calculada:

$$U^* = \frac{1}{2} \delta u^t K_5 \delta u + \frac{1}{2} \delta w^t K_6 \delta w \quad (2.40)$$

Da literatura clássica de resistência dos materiais tem-se que

$$\alpha = \frac{12 EI}{GL^2 S_r} \quad (2.41)$$

onde

$$G = \frac{E}{2(1-\nu)} \quad (2.42)$$

ν é o coeficiente de Poisson

G é o módulo de Coulomb

S_r é a secção equivalente ao cisalhamento

Então a energia potencial total é dada pela equação (2.43).

$$U_A = \frac{1}{2} \delta u^t K_5 \delta u + \frac{1}{2} \delta w^t K_6 \delta w + \frac{1}{2} \delta u^t K_3 \delta u + \frac{1}{2} \delta w^t K_4 \delta w \quad (2.43)$$

A introdução de α , o coeficiente de cisalhamento de Timoshenko, modifica as matrizes K_1 e K_2 para as matrizes K_5 e K_6 .

$$K_5 = \frac{EI}{L^3} \begin{vmatrix} \frac{12}{1+\alpha} & \frac{-6L}{1+\alpha} & \frac{-12}{1+\alpha} & \frac{-6L}{1+\alpha} \\ \frac{-6L}{1+\alpha} & \frac{4+\alpha}{4+\alpha} L^2 & \frac{6L}{1+\alpha} & \frac{2-\alpha}{2-\alpha} L^2 \\ \frac{12}{1+\alpha} & \frac{6L}{1+\alpha} & \frac{12}{1+\alpha} & \frac{6L}{1+\alpha} \\ \frac{-6L}{1+\alpha} & \frac{2-\alpha}{2-\alpha} L^2 & \frac{-6L}{1+\alpha} & \frac{4+\alpha}{4+\alpha} L^2 \end{vmatrix} \quad (2.44)$$

$$K_6 = \frac{EI}{L^3} \begin{vmatrix} \frac{12}{1+\alpha} & \frac{6L}{1+\alpha} & \frac{-12}{1+\alpha} & \frac{6L}{1+\alpha} \\ \frac{6L}{1+\alpha} & \frac{4+\alpha}{4+\alpha} L^2 & \frac{-6L}{1+\alpha} & \frac{2-\alpha}{2-\alpha} L^2 \\ \frac{12}{1+\alpha} & \frac{-6L}{1+\alpha} & \frac{12}{1+\alpha} & \frac{-6L}{1+\alpha} \\ \frac{-6L}{1+\alpha} & \frac{2-\alpha}{2-\alpha} L^2 & \frac{-6L}{1+\alpha} & \frac{4+\alpha}{4+\alpha} L^2 \end{vmatrix} \quad (2.45)$$

Aplicando a equação (2.1), equação de Lagrange, na expressão da energia potencial total, equação (2.43), obtém-se:

$$\frac{\partial U_A}{\partial q_i} = [K + K_F] \{q\} \quad (2.46)$$

onde:

- K é a matriz rigidez total devido à flexão e ao cisalhamento, obtida das matrizes K_5 e K_6 .
- K_F é a matriz rigidez total devido ao carregamento axial do elemento árvore, obtida das matrizes K_3 e K_4 .
- K e K_F são mostradas pelas equações (2.47) e (2.48).

$$K = \frac{EI}{(1+\alpha)L^3} \begin{vmatrix} 12 & 0 & 0 & -6L & -12 & 0 & 0 & -6L \\ 0 & 12 & 6L & 0 & 0 & -12 & 6L & 0 \\ 0 & 6L & (4+\alpha)L^2 & 0 & 0 & -6L & (2-\alpha)L^2 & 0 \\ -6L & 0 & 0 & (4+\alpha)L^2 & 6L & 0 & 0 & (2-\alpha)L^2 \\ -12 & 0 & 0 & 6L & 12 & 0 & 0 & 6L \\ 0 & -12 & -6L & 0 & 0 & 12 & -6L & 0 \\ 0 & 6L & (2-\alpha)L^2 & 0 & 0 & -6L & (4+\alpha)L^2 & 0 \\ -6L & 0 & 0 & (2-\alpha)L^2 & 6L & 0 & 0 & (4+\alpha)L^2 \end{vmatrix} \quad (2.47)$$

$$K_F = \frac{F_0}{30L} \begin{bmatrix} 36 & 0 & 0 & -3L & -36 & 0 & 0 & -3L \\ 0 & 36 & 3L & 0 & 0 & -36 & 3L & 0 \\ 0 & 3L & 4L^2 & 0 & 0 & -3L & -L^2 & 0 \\ -3L & 0 & 0 & 4L^2 & 3L & 0 & 0 & -L^2 \\ -36 & 0 & 0 & 3L & 36 & 0 & 0 & 3L \\ 0 & -36 & -3L & 0 & 0 & 36 & -3L & 0 \\ 0 & 3L & -L^2 & 0 & 0 & -3L & 4L^2 & 0 \\ -3L & 0 & 0 & -L^2 & 3L & 0 & 0 & 4L^2 \end{bmatrix} \quad (2.48)$$

2.4 - Mancais.

Mancais são dispositivos responsáveis pela ligação entre a parte móvel e a estrutura fixa de uma máquina rotativa. Um modelo físico simplificado permite representá-lo pelas seguintes características principais:

- a inércia polar,
- a inércia transversal,
- a rigidez e
- o amortecimento.

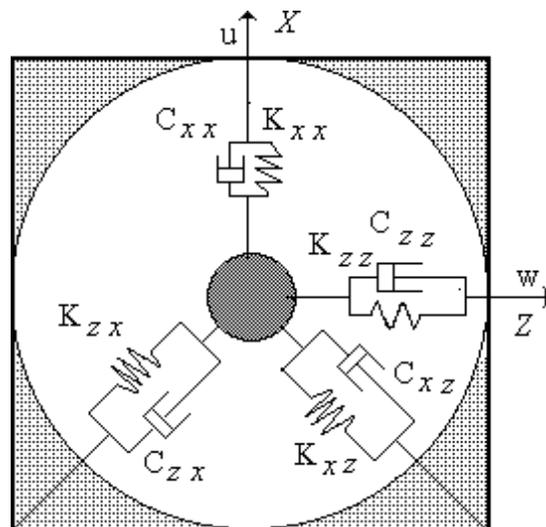


Figura (2.5) - Modelo de um mancal

K_{xx} , C_{xx} e K_{yy} , C_{yy} são respectivamente rigidez e amortecimento nas direções do eixo X e do eixo Z .

K_{xz} e C_{xz} são rigidez e amortecimento aplicados na direção X porém com efeitos na direção do eixo Z .

A figura (2.5) mostra graficamente a configuração de um mancal. Além de absorvedor de energia, outra importante função dos mancais é o de guiar ou restringir graus de liberdade durante a movimentação da árvore. Considerando apenas o comportamento linear do mancal, as forças exercidas pelo mancal sobre a árvore são F_u e F_w .

A equação (2.49) mostra a expressão do trabalho virtual δw realizados pelas forças que atuam sobre o mancal.

$$\delta w = F_u \delta u + F_w \delta w \quad (2.49)$$

Ou:

$$\begin{aligned} \delta w = & -k_{xx} u \delta u - k_{xz} w \delta u - k_{zz} w \delta w - k_{zx} u \delta w - \\ & - c_{xx} \dot{u} \delta u - c_{xz} \dot{w} \delta u - c_{zz} \dot{w} \delta w - c_{zx} \dot{u} \delta w \end{aligned} \quad (2.50)$$

$$\begin{aligned} F_u = & -K_{xx} u - K_{xz} w - C_{xx} \dot{u} - C_{xz} \dot{w} \\ F_w = & -K_{zx} u - K_{zz} w - C_{zx} \dot{u} - C_{zz} \dot{w} \end{aligned} \quad (2.51)$$

Na forma matricial:

$$\begin{Bmatrix} F_u \\ F_w \end{Bmatrix} = \begin{Bmatrix} K_{xx} & K_{xz} \\ K_{zx} & K_{zz} \end{Bmatrix} \begin{Bmatrix} u \\ w \end{Bmatrix} - \begin{Bmatrix} C_{xx} & C_{xz} \\ C_{zx} & C_{zz} \end{Bmatrix} \begin{Bmatrix} \dot{u} \\ \dot{w} \end{Bmatrix} \quad (2.52)$$

2.5 - Excitação por desbalanceamento

Excitação por desbalanceamento é a fonte de sollicitação mais comumente encontrada nas máquinas rotativas. A causa do desbalanceamento é assumido como sendo o efeito produzido por massas excêntricas localizadas em planos e em posições angulares diferentes ao logo do rotor, como mostra a figura (2.6).

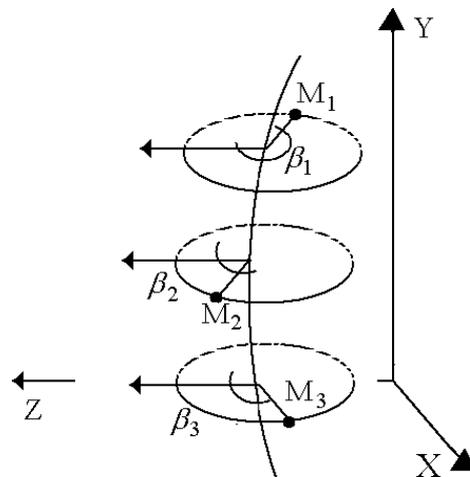


Figura (2.6) - Excitação por desbalanceamento

A posição M de cada massa m é definida pela:

- sua abscissa l , ao longo do eixo Y ,
- a distancia $a = O_1M$, onde O_1 é centro da secção no plano considerado,
- pelo ângulo $\beta = \langle Z, O_1M \rangle$.

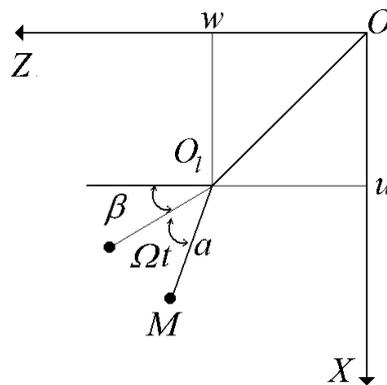


Figura (2.7) - Massa de desbalanceamento.

Seja o desbalanceamento num dado plano, como mostra a figura (2.7). A posição da massa de desbalanceamento, escrita no referencial inercial é dada pela equação (2.53) onde:

$$\vec{OM} = \vec{OO_1} + \vec{O_1M} \tag{2.53}$$

ou

$$\vec{OM} = \begin{Bmatrix} u + d \text{sen}(\Omega t + \beta) \\ y \\ w + d \text{cos}(\Omega t + \beta) \end{Bmatrix} \tag{2.54}$$

A velocidade da massa de desbalanceamento é obtida derivando a equação anterior.

$$\vec{V} = \frac{d}{dt} \vec{OM} = \begin{Bmatrix} \dot{u} + d\Omega \cos(\Omega t + \beta) \\ 0 \\ \dot{w} - d\Omega \sin(\Omega t + \beta) \end{Bmatrix} \quad (2.55)$$

A energia cinética da massa de desbalanceamento T_m , é assim definida:

$$T_m = \frac{1}{2} m \left((\dot{u} + d\Omega \cos(\Omega t + \beta))^2 + (\dot{w} - d\Omega \sin(\Omega t + \beta))^2 \right) \quad (2.56)$$

Aplicando, as equações de Lagrange, na equação (2.56) obtém-se as forças resultantes devido à presença do desbalanceamento.

$$\frac{d}{dt} \left(\frac{\partial T_m}{\partial \dot{q}_i} \right) - \frac{\partial T_m}{\partial q_i} = \begin{bmatrix} m & 0 \\ 0 & m \end{bmatrix} \begin{Bmatrix} \ddot{u} \\ \ddot{w} \end{Bmatrix} - \Omega^2 \sin \Omega t \begin{Bmatrix} \cos \beta \sin \Omega t + \sin \beta \cos \Omega t \\ \sin \beta \sin \Omega t + \cos \beta \cos \Omega t \end{Bmatrix} \quad (2.57)$$

ou

$$\begin{Bmatrix} F_u \\ F_w \end{Bmatrix} = \begin{Bmatrix} ma \cos \beta \\ -ma \sin \beta \end{Bmatrix} \Omega^2 \sin \Omega t + \begin{Bmatrix} ma \sin \beta \\ ma \cos \beta \end{Bmatrix} \Omega^2 \cos \Omega t \quad (2.58)$$

ou ainda

$$\begin{aligned} F_u &= ma \Omega^2 \sin(\Omega t + \beta) \\ F_w &= ma \Omega^2 \cos(\Omega t + \beta) \end{aligned} \quad (2.59)$$

Para forças assíncronas de amplitude F , a formulação apresentada pelas equações (2.59) permanecem válidas fazendo $\Omega = \varphi$, onde φ é diferente da frequência de rotação, logo:

$$\begin{aligned} F_{ass} &= F_u + F_w \\ F_u &= F \sin(\varphi t + \beta) \\ F_w &= F \cos(\varphi t + \beta) \end{aligned} \quad (2.60)$$

Para excitações por forças harmônicas de amplitude F e de frequência φ , fixa no espaço, excita o sistema em apenas uma direção, logo as equações (2.60) tornam-se:

$$\begin{aligned} F_{fixa} &= F_u \\ F_u &= F \sin(\varphi t + \beta) \end{aligned} \quad (2.61)$$

na direção X ou

$$\begin{aligned} F_{fixa} &= F_w \\ F_w &= F \cos(\varphi t + \beta) \end{aligned} \quad (2.62)$$

na direção Z.

Considerando o caso no qual o eixo Y do rotor, está na vertical, a força peso dos elementos de disco e árvore não exerce influência sobre a flexão do rotor.

2.6 - Equações de Movimento do Rotor

Se $q = \{x \mid y \mid \theta \mid \Psi\}^T$ é o vetor dos deslocamentos nodais, então o sistema de equações diferenciais que descreve o movimento do rotor é dado por:

$$M\ddot{q} + (C + G)\dot{q} + Kq = F_{exc} \quad (2.63)$$

onde:

$$F_{exc} = F \sin \Omega t + F \cos \Omega t \quad (2.64)$$

- M é a matriz de massa global do sistema,
- C é a matriz amortecimento oriunda dos mancais,
- G é a matriz de efeito giroscópico do rotor,
- K é a matriz das rigidez oriunda do mancais e dos elementos de arvores,
- F_{exc} é o vetor das forças de excitações,
- Ω é a frequência da força de excitação.

A equação (2.64) representa as forças harmônicas de desbalanceamento, assíncronas ou forças fixas no espaço.

2.7 - Resposta de Sistema Rotativo à Excitações: Método Pseudo Modal

A equação (2.63) representa uma equação matricial diferencial de segunda ordem não homogênea. Classicamente, a solução total da equação (2.63) é dada pela soma da solução da equação homogênea associada mais a solução particular, uma vez que o sistema é linear.

O **método pseudo modal** (Engels, R. C., 1991) é usado para reduzir a ordem de grandeza da equação (2.63) utilizando apenas alguns dos primeiros modos do sistema não giroscópico associado. Este método permite reduzir o número de graus de liberdade do sistema original, simplificando a sua solução e diminuindo substancialmente o esforço computacional necessário à sua obtenção.

A resposta do sistema às excitações síncronas ou assíncronas, bem como os modo de freqüências são calculados para o sistema reduzido sem substancial prejuízo da precisão numérica (Engels, R. C., 1991).

Seja o sistema homogêneo ($F_{exc} = 0$), não giroscópico ($\Omega = 0$) e não amortecido ($C = 0$) associado ao sistema da equação (2.63):

$$M\ddot{q} + Kq = 0 \quad (2.65)$$

Seja a solução da equação (2.65), do tipo:

$$q = q_0 e^{pt} \quad (2.66)$$

então:

$$p^2 M q_0 + K q_0 = 0 \quad (2.67)$$

$$K^{-1} M q_0 = -\frac{1}{p^2} q_0 \quad (2.68)$$

Da equação (2.68) obtém-se os primeiros p_n autovalores e os correspondentes $\Phi = [\Phi_1, \Phi_2, \dots, \Phi_n]$ autovetores.

Premultiplicando todos os termos da equação (2.63) por Φ^t e fazendo a seguinte mudança de coordenadas:

$$\Phi q = \Delta \quad (2.69)$$

tem-se que:

$$\Phi^t M \Phi \ddot{q} + \Phi^t (C + G) \Phi \dot{q} + \Phi^t K \Phi q = \Phi^t F_{exc} \quad (2.70)$$

ou

$$M^* \Delta + C^* \Delta + K^* \Delta = F_{exc}^* \quad (2.71)$$

onde

$$\begin{aligned} M^* &= \Phi^t M \Phi \\ C^* &= \Phi^t (C + G) \Phi \\ K^* &= \Phi^t K \Phi \\ F_{exc}^* &= \Phi^t F_{exc} \end{aligned} \quad (2.72)$$

A equação (2.71) é uma versão reduzida da equação (2.63).

2.7.1 - Solução da equação homogênea

Da equação (2.71) tem-se que:

$$M^* \ddot{\Delta} + C^* \dot{\Delta} + K^* \Delta = 0 \quad (2.73)$$

Combinando a equação anterior com

$$-M^* \dot{\Delta} + M^* \dot{\Delta} = 0 \quad (2.74)$$

obtem-se:

$$\left[\begin{array}{c|c} 0 & -M^* \\ \hline M^* & C^* \end{array} \right] \begin{Bmatrix} \ddot{\Delta} \\ \dot{\Delta} \end{Bmatrix} + \left[\begin{array}{c|c} M^* & 0 \\ \hline 0 & K^* \end{array} \right] \begin{Bmatrix} \dot{\Delta} \\ \Delta \end{Bmatrix} = 0 \quad (2.75)$$

ou

$$A \dot{s} + B s = 0 \quad (2.76)$$

Seja a solução do tipo:

$$s = s_0 e^{r t} \quad (2.77)$$

logo a equação (2.63) torna-se:

$$B^{-1} A s_0 = -\frac{1}{r} s_0 \quad (2.78)$$

ou

$$D s_0 = \lambda s_0 \quad (2.79)$$

A equação (2.79) permite, através de métodos numéricos, calcular os $2n$ autovalores e autovetores complexos que facilmente são convertidos às coordenadas originais através da transformação $\Delta = \Phi q$.

2.7.2 - Solução Particular: Resposta à Excitações

Considerando a linearidade da equação de movimento do rotor, equação (2.71), a propriedade da superposição de efeitos é válida. Na dinâmica de rotação, como dito anteriormente, as excitações harmônicas clássicas são as advindas de forças de desbalanceamento, forças assíncronas e forças fixas no espaço.

$$M^* \Delta + C^* \dot{\Delta} + K^* \Delta = F_1 \sin \omega t + F_2 \cos \omega t \quad (2.78)$$

Para forças de desbalanceamento $F_1 = F_2 = ma \Omega^2$.

Para forças assíncronas $\Omega \neq \omega$.

Para forças fixa no espaço $F_1 = 0$ ou $F_2 = 0$, dependendo da orientação espacial.

Para a solução particular procura-se solução do tipo

$$\Delta = \delta_1 \sin \omega t + \delta_2 \cos \omega t \quad (2.81)$$

Substituindo a equação (2.81) na equação (2.80) e isolando os termos afins, tem-se que:

$$\begin{bmatrix} -\omega^2 M^* + K^* & -C^* \\ C^* & -\omega^2 M^* + K^* \end{bmatrix} \begin{Bmatrix} \delta_1 \\ \delta_2 \end{Bmatrix} = \begin{Bmatrix} F_1 \\ F_2 \end{Bmatrix} \quad (2.82)$$

Após a solução da sistema anterior, o retorno às coordenadas originais se dá através da transformação:

$$\Delta = \Phi q \quad (2.83)$$

Capítulo 3

Introdução à Análise de Sinal

3.1 - Introdução

Vários trabalhos têm demonstrado a importância e a potencialidade de cada uma das várias ferramentas matemáticas, agrupadas sob o rótulo de análise de sinais, na solução de problemas mecânicos ou auxiliando para sua compreensão.

Fregolent (Fregolent et al, 1996) apresenta uma formulação baseada na teoria de transformada de wavelet. Este trabalho mostra a determinação de um conjunto de forças que atuam numa estrutura, usando o sinal de resposta do sistema. Esta formulação, inicialmente proposta por Euler-Bernoulli, é capaz de identificar a localização e a amplitude das forças aplicadas numa estrutura, mesmo quando aplicadas em pontos inacessíveis à medição. Este procedimento usa as equações de movimento da estrutura. Os parâmetros geométricos e físicos devem ser conhecidos. Não é necessário o conhecimento a priori das condições de contorno, uma vez que elas possam ser experimentalmente mensuradas. As informações de posição e amplitude das forças aplicadas são extraídas do espectro de forças no domínio de wavelet, inspecionando os componentes harmônicas.

O filtro de Kalman, inicialmente proposto (Kalman, 1960) para estimar estados não conhecidos de um sistema linear, a partir de medidas de vibração, foi ampliado (Jazwinski, 1970) para aplicação em sistemas não-lineares.

O uso de um banco de filtros de Kalman (Seibold, 1996), cada um condicionado para uma específica condição de um dado sistema, tem sido proposto para a indicação de uma dada ocorrência. Essa técnica consiste em otimizar um algoritmo de múltipla hipóteses com respeito as diferentes falhas e seus efeitos dinâmicos sobre o sistema mecânico, tendo como base um algoritmos de quantificação de confiança que leva em conta: 1 - o modelo mecânico do sistema; 2 - um critério de falha, especificando as condições que se constitui uma falha ou um algoritmos de diagnostico e, 3 - um modelo de incerteza, quantificando as incertezas que o modelo do sistema está sujeito, quanto ao carregamento bem como do critério de falha.

Varias publicações (Lembregts et al, 1996) (Romberg et al, 1996) descrevem os fundamentos e o uso de técnicas de processamento de sinal bem estabelecidas, como amostragem digital de sinais e filtros de conversão de sinais de um domínio em outro, remoção de ruído, análise de dados quase-estacionários, remoção de flutuação de rotação, sincronização de médias, compensação do efeito doppler, etc.. Em particular, ilustra como determinados fenômenos podem ser mais facilmente tratados num determinado domínio de que em outros.

Método baseado na densidade da potência espectral (Nauerz et al, 1996) tem sido aplicado para a localização da variação de rigidez num modelo de elementos finitos. Este método é aplicável para o caso em que as excitações não são conhecidas, mas tendem para um processo aleatório estacionário, cujas propriedades estatísticas como média e covariância pode ser estimada. Usando análise de sensibilidade, é montado um conjunto de equações para o sistema linear associado, que são resolvidas usando medidas de resposta experimentais, tratados por técnicas de filtro de Kalman.

A transformada de wavelet (Ruzzene, 1997) tem sido usada para representar sistemas nos domínios do tempo e da frequência, nos processos de identificação dos mais diversos parâmetros tais como frequências naturais e fator de amortecimento viscoso.

3.2 - Revisão de Análise de Sinal

Um processo $\{x_k(t)\}$ para $-\infty < t < \infty$ é dito processo aleatório quando ele não pode ser descrito por uma relação matemática explícita. Cada sinal $x(t)$ amostrado no tempo é único e representa um dos possíveis resultados do sistema.

Um sinal aleatório somente pode ser caracterizado através de suas propriedades estatísticas. Além disso, um sinal aleatório é chamado estacionário se suas propriedades estatísticas são invariantes com respeito ao tempo (Romberg, T. M., 1996).

As propriedades estatísticas de importância na descrição de um único sinal aleatório são:

- valores de média e média quadrática
- funções de densidade de probabilidade
- função auto correlação
- função densidade espectral

O valor da média μ_x e da variância σ_x^2 para um sinal estacionário, representa respectivamente a tendência central e a dispersão dos dados.

A média quadrática ψ_x^2 que é dado pela variância mais o quadrado da média, constitui-se numa medida que combina tendência central e dispersão.

Para pares de sinais aleatórios amostrados de diferentes processos as propriedades estatísticas conjuntas de importância são:

- função densidade de probabilidade conjunta
- função correlação cruzada
- função densidade espectral cruzada
- função resposta em frequência
- função coerência

A função de densidade de probabilidade permite:

- a indicação de efeitos não lineares e
- análise de valores extremos.

As medidas de correlação permitem:

- detecção de periodicidades, medidas de intervalo de tempo,
- análise de ruído,
- propagação de trajetórias e velocidades.
- a determinação das propriedades do sistema a partir dos dados de saída ou de entrada,
- a obtenção da resposta do sistema a partir de suas características de dos dados de entrada,
- identificação de energias e fontes de ruídos, etc.

3.3 - Análise no Domínio do Tempo

Dado $x_k(t)$, um conjunto de amostras, pode-se estimar um sinal médio através da seguinte equação:

$$\bar{x}(t) = \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{i=0}^N x_i(t_1) \quad (3.1)$$

Se o sinal médio não variar para o conjunto de amostras $x_k(t)$, colhidas em tempos arbitrários, o sinal é dito estacionário ou estocástico. Por outro lado, define-se valor médio do sinal como:

$$\mu_x = \lim_{N \rightarrow \infty} \frac{1}{T} \int_0^T x(t) dt \quad (3.2)$$

Se o valor médio μ_x calculado para diferentes $x(t)$ for invariante, o sinal $x(t)$ é chamado ergótico.

A média quadrática é assim definida equação (3.3).

$$\Psi_x^2 = \lim_{N \rightarrow \infty} \frac{1}{T} \int_0^T x^2(t) dt \quad (3.3)$$

O valor da média μ_x e da variância σ_x^2 para um sinal estacionário, representa respectivamente a tendência central e a dispersão dos dados.

A média quadrática ψ_x^2 que é dado pela variância mais o quadrado da média, e constitui-se numa medida que combina tendência central e dispersão.

Funções de correlação são uma extensão dos conceitos de sinal médio e valor médio do sinal, aplicadas para amostras defasadas por um intervalo de tempo arbitrário.

A função autocorrelação é definida como:

$$R_{xx}(t_1, \lambda) = \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{i=1}^N x_i(t_1) x_i(t_1 + \lambda) \quad (3.4)$$

onde λ é a defasagem entre a amostragem dos sinais.

Se o processo é ergótico, qualquer amostra é representativa do sinal médio logo:

$$R_{xx}(t, \lambda) = \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T x(t) x(t + \lambda) dt \quad (3.4)$$

Considerando nula a defasagem entre os sinais tem-se que:

$$\begin{aligned} R_{xx}(0) &= \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T x^2(t) dt \\ C_{yy}(\lambda) &= R_{xx}(\lambda) - \mu_y^2 \\ C_{xy}(\lambda) &= R_{xx}(\lambda) - \mu_y^2 \end{aligned} \quad (3.5)$$

Normalizando a função de autocorrelação pelo valor médio quadrático $R_{xx}(0)$ obtém-se a função autocorrelação normalizada.

$$\Phi(\lambda) = \frac{R_{xx}(\lambda)}{R_{xx}(0)} \quad (3.6)$$

Removendo o quadrado do valor médio das funções de autocorrelação, obtém-se as funções de autocovariância, dadas pelas equações seguintes:

$$\begin{aligned} C_{xx}(\lambda) &= R_{xx}(\lambda) - \mu_x^2 \\ C_{yy}(\lambda) &= R_{yy}(\lambda) - \mu_y^2 \\ C_{xy}(\lambda) &= R_{xy}(\lambda) - \mu_x \mu_y \end{aligned} \quad (3.7)$$

Estendendo esses conceitos para sinais de variáveis diferentes, obtém-se as funções de correlações cruzadas:

$$\begin{aligned} R_{xy}(\lambda) &= \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T x(t) y(t + \lambda) dt \\ R_{yx}(\lambda) &= \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T x(t + \lambda) y(t) dt \\ \Phi(\lambda) &= \frac{R_{xy}(\lambda)}{\sqrt{R_{xx}(0) R_{yy}(0)}} \end{aligned} \quad (3.8)$$

3.4 - Análise no domínio da frequência

Na seção anterior, viu-se que um sinal aleatório pode ser caracterizado no domínio do tempo, através do uso das funções de autocorrelação; e que dois sinais distintos podem ser analisados conjuntamente através das funções de correlação cruzadas. Porém, em algumas aplicações, é vantajoso trabalhar no domínio da frequência, pois a visualização das propriedades do sinal é direta e intuitiva.

3.4.1 - Função densidade de potência

Assumindo que um sinal seja periódico, o mesmo pode ser representado em Série de Fourier, ou seja:

$$x(t) = \sum_{n=-\infty}^{\infty} C_n e^{jn\omega_0 t} \quad (3.9)$$

onde:

$$C_n = \frac{1}{T} \int_0^T x(t) e^{-jn\omega_0 t} dt$$

$$\omega_0 = \frac{2\pi}{T} \quad (3.10)$$

A potência do n -ésimo componente é simétrico com relação à origem e é dado por:

$$C_n^2 = \frac{C_{-n}^2}{2} + \frac{C_n^2}{2} \quad (3.11)$$

Considerando apenas um lado da simetria, define-se a potência por unidade de banda ou

$$\text{densidade de potência} = \frac{C_n^2}{1/T} \quad (3.12)$$

cuja unidade é a relação amplitude ao quadrado por Hertz.

Para sinais aleatórios não periódicos, a Série de Fourier não pode ser aplicada diretamente. A estratégia empregada nesses casos consiste em gerar artificialmente, sinais periódicos $x_p(t)$, a partir do processo aleatório, removendo assim o impedimento para o uso da equação (3.9) que pode ser reescrita como:

$$x_p(t) = \sum_{n=-\infty}^{\infty} C_n e^{jn\omega_0 t} \quad (3.13)$$

onde:

$$C_n = \frac{1}{T} \int_{-T/2}^{+T/2} x_p(t) e^{-jn\omega_0 t} dt$$

$$\omega_0 = \frac{2\pi}{T} \quad (3.14)$$

fazendo $X_n = C_n T$, tem-se o par:

$$X_n = \int_{-T/2}^{+T/2} x_p(t) e^{-jn\omega_0 t} dt$$

$$x_p(t) = \sum_{n=-\infty}^{\infty} X_n e^{jn\omega_0 t} \frac{1}{T} \quad (3.15)$$

da seção anterior tem-se então que a função densidade espectral pode ser representada pela equação 3.16.

$$\text{Densidade de potência} = \frac{1}{T} X_n^2 \quad (3.16)$$

Fazendo as seguintes considerações:

$$\frac{1}{T} \rightarrow df$$

$$x_p(t) \rightarrow x(t) \quad (3.17)$$

$$n\omega_0 \rightarrow \omega$$

$$X_n \rightarrow X(\omega)$$

o par de transformada de Fourier pode ser reescrito como:

$$X(\omega) = \int_{-\infty}^{+\infty} x(t) e^{j\omega t} dt$$

$$x(t) = \int_{-\infty}^{+\infty} X(\omega) e^{-j\omega t} dt \quad (3.18)$$

a densidade de potência, equação (3.16), torna-se

$$\frac{X_n^2}{T} \rightarrow \frac{1}{T} X(\omega) X^*(\omega) \quad (3.19)$$

onde * indica o complexo conjugado.

A densidade de potência espectral é definido como:

$$S_{xx}(\omega) = \lim_{T \rightarrow \infty} \frac{1}{T} \{X(\omega) X^*(\omega)\} \quad (3.20)$$

com $-\infty < \omega < +\infty$.

A densidade de potência mensurável é definida como

$$G_{xx}(\omega) = 2S_{xx}(\omega) \quad (3.21)$$

onde $0 < \omega < +\infty$. A figura (3.1) mostra a relação entre essas duas grandezas.

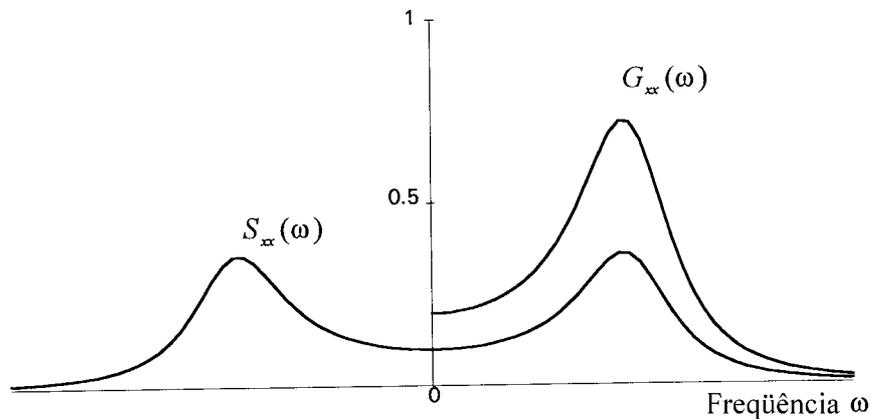


Figura (3.1) - Densidade de potência teórica e densidade de potência mensurável.

As densidade espectral cruzadas calculadas entre dois sinais são dadas por:

$$\begin{aligned} S_{xy}(\omega) &= \lim_{T \rightarrow \infty} \frac{1}{T} \{X^*(\omega) Y(\omega)\} \\ S_{yx}(\omega) &= \lim_{T \rightarrow \infty} \frac{1}{T} \{Y^*(\omega) X(\omega)\} \end{aligned} \quad (3.22)$$

3.4.2 - Transformada Rápida de Fourier (FFT)

A Transformada de Rápida de Fourier é uma ferramenta computacional eficiente, amplamente usada nos problemas de manipulação de dados mensurados.

Um processo físico pode ser descrito no domínio do tempo por um conjunto discreto de valores de $h(t)$, ou no domínio da frequência, pelo conjunto $H(f)$, geralmente um número complexo que traz a informação de fase, tal que $-\infty < f < \infty$. $H(f)$ e $h(t)$ são duas representações de um mesmo fenômeno cuja transformada de Fourier é dada por:

$$\begin{aligned}
 h(t) &= \int_{-\infty}^{\infty} H(f) e^{-2\pi i f t} df \\
 H(f) &= \int_{-\infty}^{\infty} h(t) e^{2\pi i f t} dt
 \end{aligned}
 \tag{3.23}$$

onde t é medido em segundos e f é dado em ciclos por segundo. Lembrado que a frequência angular é dado em radianos por segundo tal que $\omega = 2\pi f$, a equação (3.23) torna-se:

$$\begin{aligned}
 h(t) &= \frac{1}{2\pi} \int_{-\infty}^{\infty} H(\omega) e^{-i\omega t} d\omega \\
 H(f) &= \int_{-\infty}^{\infty} h(t) e^{i\omega t} dt
 \end{aligned}
 \tag{3.24}$$

A seguir tem-se um resumo das simetrias existentes entre o dois domínios, $h(t)$ e $H(f)$.

$h(t)$ é real	\Rightarrow	$H(-f) = H(f) ^*$
$h(t)$ é imaginário	\Rightarrow	$H(-f) = - H(f) ^*$
$h(t)$ é par	\Rightarrow	$H(-f) = H(f)$
$h(t)$ é ímpar	\Rightarrow	$H(-f) = -H(f)$
$h(t)$ é real e par	\Rightarrow	$H(f)$ é real e par
$h(t)$ é real e ímpar	\Rightarrow	$H(f)$ é real e ímpar
$h(t)$ é imaginário e par	\Rightarrow	$H(f)$ é imaginário e par
$h(t)$ é imaginário e ímpar	\Rightarrow	$H(f)$ é imaginário e ímpar

Segue-se outras propriedades elementares da Transformada de Fourier. Se $h(t) \Leftrightarrow H(f)$, então:

$$h(at) \Leftrightarrow \frac{1}{|a|} H(f/a) \quad \text{escalonamento do tempo} \tag{3.25}$$

$$\frac{1}{|b|} h(t/b) \Leftrightarrow H(bf) \quad \text{escalonamento da frequência} \tag{3.26}$$

$$h(t-t_0) \Leftrightarrow H(f) e^{2\pi i f t_0} \quad \text{deslocamento no tempo} \tag{3.27}$$

$$h(t) e^{-2\pi i f_0 t} \Leftrightarrow H(f-f_0) \quad \text{deslocamento na frequência} \tag{3.28}$$

Seja $h(t)$ e $g(t)$ duas funções cujas correspondentes Transformada de Fourier são $H(f)$ e $G(f)$. Algumas combinações de $h(t)$ e $g(t)$ são especialmente interessantes.

A convolução das duas funções, no domínio do tempo, é dada por:

$$g * h \equiv \int_{-\infty}^{\infty} g(\tau) h(t - \tau) d\tau \quad (3.29)$$

Pelo teorema da convolução tem-se que

$$g * h \Leftrightarrow G(f)H(f) \quad (3.30)$$

A correlação de duas funções, no domínio do tempo é dada por

$$R_{g,h}(t) \equiv \int_{-\infty}^{\infty} g(\tau + t) h(\tau) d\tau \quad (3.31)$$

No domínio da frequência tem-se que

$$R_{g,h}(t) \Leftrightarrow G(f)H^*(f) \quad (3.32)$$

A potência total de um sinal pode ser calculada tanto nos domínios do tempo e da frequência. Este resultado é conhecido como teorema de Parseval:

$$\text{Potência Total} \equiv \int_{-\infty}^{\infty} |h(t)|^2 dt = \int_{-\infty}^{\infty} |H(f)|^2 df \quad (3.33)$$

3.4.3 - Transformada de Fourier de dados discretos

As situações mais comuns, a função $h(t)$ é apenas conhecida (amostrada) em intervalos regulares de tempo, onde Δ denota o tempo de espaçamento, tal que:

$$h_n = h(n\Delta) \quad n = \dots, -2, -1, 0, 1, 2, \dots \quad (3.34)$$

$1/\Delta$ é a taxa de amostragem. Para cada intervalo de amostragem Δ , existe uma frequência especial f_c , chamada frequência crítica de Nyquist, dada por:

$$f_c \equiv \frac{1}{2\Delta} \quad (3.35)$$

^{3.1}Largura de banda é o intervalo que contém todas as frequências presentes num sinal.

ou seja, a amostragem mínima de um a onda senoidal é dois pontos por ciclo.

Teorema da amostragem

Se uma função contínua $h(t)$, for amostrada em intervalos de tempo Δ , ela possui largura de banda^{3.1} limitada em magnitude, para freqüências menores que f_c , isto é, $H(f) = 0$ para todo $|f| \geq f_c$; conseqüentemente, $h(t)$ é completamente determinada pelas suas amostras h_n . Logo:

$$h(t) = \Delta \sum_{n=-\infty}^{+\infty} h_n \frac{\sin[2\pi f_c(t - n\Delta)]}{\pi(t - n\Delta)} \quad (3.36)$$

Na prática, a amostragem das funções contínuas não tem largura de banda limitada para freqüências menores que a freqüência crítica de Nyquist ($-f_c < f < f_c$). Todas as freqüências situadas fora dessa faixa são falsamente transladadas para dentro da faixa. Este fenômeno é chamado de aliasing. A figura (3.2) mostra graficamente o efeito desse fenômeno. Este efeito somente poderá ser eliminado através da aplicação de filtro passa-baixa, no sinal da função original antes de ser amostrada.

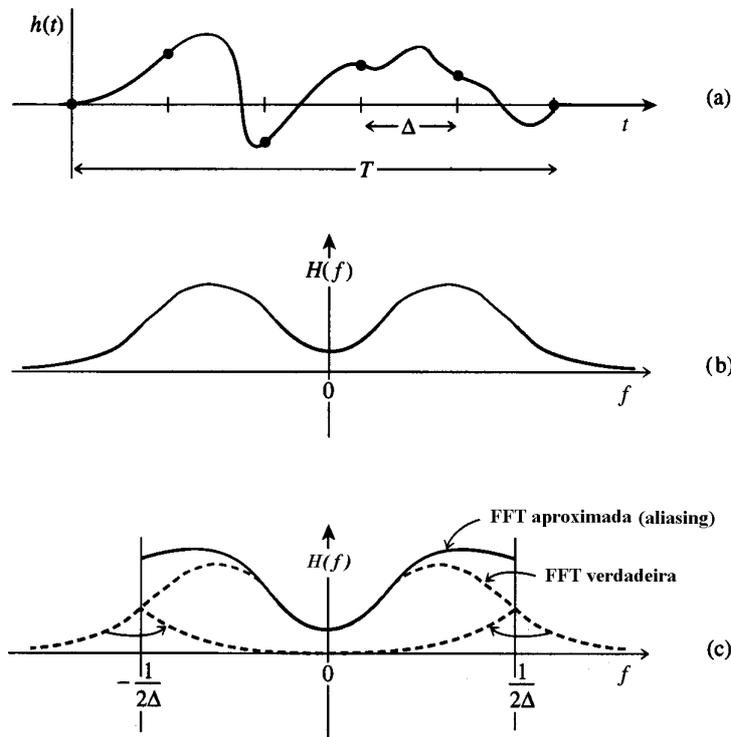


Figura (3.2) - Fenômeno aliasing. (a) Sinal original. (b) FFT do sinal original. (c) Efeito do aliasing

3.4.4 - Correlação e Autocorrelação usando FFT

A correlação é uma operação matemática semelhante à convolução. A diferença é que, duas funções que serão correlacionadas não são conceitualmente tão distintas quanto os dados e funções respostas participantes de uma convolução.

A correlação é uma operação realizada entre conjuntos de dados diferentes que representativos um mesmo processo, verificando o grau de superposição, de deslocamento à direita ou à esquerda, entre as amostras comparadas.

Como definido pela equação (3.31), a correlação entre duas funções contínuas $g(t)$ e $h(t)$, que é denotada por $R_{g,h}$ ou $R_{g,h}(t)$, e é uma função defasada com relação a t , dependência esta que pode ser mostrada pela correlação $R_{g,h}(t)$.

A correlação $R_{g,h}(t)$ assume valores de grande magnitude para alguns valores de positivos de t , se a função $g(t)$ for uma cópia da função $h(t)$ defasada do tempo t , ou seja, $g(t)$ está deslocada à direita de $h(t)$. Por outro lado, $R_{g,h}(t)$ assume valores de grande magnitude para valores de negativos de t se $g(t)$ está deslocada à esquerda de $h(t)$, logo

$$R_{g,h}(t) = R_{h,g}(-t) \quad (3.37)$$

A correlação para duas funções discretas g_k e h_k de período N , é definida por

$$R_{g,h} \Big|_j = \sum_{k=0}^{N-1} g_{j+k} h_k \quad (3.38)$$

A teoria da correlação discreta (Nussbaumer. H. J. 1982) diz que a correlação discreta de duas funções reais g e h é um membro do par de transformada discreta da Fourier

$$R_{g,h} \Big|_j \Leftrightarrow G_k H_k^* \quad (3.39)$$

onde * indica o complexo conjugado. G_k e H_k são as transformadas discretas de Fourier das funções g_j e h_j respectivamente.

O cálculo da correlação usando FFT é feito da seguinte maneira: calcula-se a FFT dos dois conjuntos de dados, multiplica um pelo complexo conjugado do outro, e calcula-se a transformada inversa de Fourier do produto. O resultado r_k é formalmente um vetor complexo com parte imaginária nula. Os valores reais de r_k são as correlações para diferentes defasagens, armazenadas na seguinte seqüência: defasagem = 0 está em r_0 , defasagem = 1 está em r_1 , etc.

3.4.5 - Densidade Espectral de Potência (PSD) estimada pela FFT

Em geral existe alguma relação de proporcionalidade entre a medida do quadrado da amplitude de uma função e a medida da amplitude da PSD. Seja a função $b(t)$ conhecida em N pontos tal que $b(t) = \{b_0, b_1, \dots, b_{N-1}\}$, e que esses pontos sejam amostrados num intervalo de tempo T tal que $T = (N-1)\Delta$, onde Δ é o intervalo de amostragem. Existem diferentes descritores da potência total:

$$\sum_{j=0}^{N-1} |b_j|^2 \equiv \text{somatório do quadrado das amplitudes} \quad (3.40)$$

$$\frac{1}{T} \int_0^T |b(t)|^2 dt \approx \frac{1}{N} \sum_{j=0}^{N-1} |b_j|^2 \equiv \text{mé dia do quadrado das amplitudes} \quad (3.41)$$

$$\int_0^T |b(t)|^2 dt \approx \Delta \sum_{j=0}^{N-1} |b_j|^2 \equiv \text{integral do quadrado das amplitudes} \quad (3.42)$$

Este estimador da PSD é uma versão simplificada do estimador chamado periodograma. Seja a FFT da função $c(t)$:

$$B_k = \sum_{j=0}^{N-1} b_j e^{2\pi i j k / N} \quad k = 0, 1, \dots, N-1 \quad (3.43)$$

então o periodograma estimador da PSD é definida em $N/2 + 1$ frequências tais que

$$\begin{aligned} P(0) &= P(f_0) = \frac{1}{N^2} |B_0|^2 \\ P(f_k) &= \frac{1}{N^2} [|B_k|^2 + |B_{N-k}|^2] \quad k = 1, 2, \dots, N/2 + 1 \\ P(f_c) &= P(f_{N/2}) = \frac{1}{N^2} |B_{N/2}|^2 \end{aligned} \quad (3.44)$$

onde f_k é definida apenas para frequências maior ou igual a zero.

$$f_k = \frac{k}{N\Delta} = 2f_c \frac{k}{N} \quad k = 0, 1, \dots, N/2 \quad (3.45)$$

A questão a ser respondida é; a equação (3.44) é um estimador “verdadeiro” da PSD da função $b(t)$, ou seja, o valor esperado do estimador periodograma é igual à densidade de potência

espectral? Na realidade não se pode esperar que o valor de $P(f_k)$ discreto seja igual ao $P(f)$ contínuo exatamente em f_k . Espera-se que $P(f_k)$ seja uma espécie de média de $P(f)$ sobre uma faixa (janela) estreita da função, centrada em f_k . Para o estimador periodograma, equação (3.44), define-se a função janela, como função de s , a freqüência de offset.

$$W(s) = \frac{1}{N^2} \left[\frac{\sin(\pi s)}{\sin(\pi s / N)} \right]^2 \quad (3.46)$$

Nota-se que a equação (3.46) possui picos oscilação que decaem não instantaneamente, para valores quase nulos, da ordem de $W(s) \approx (\pi s)^{-2}$. Este resultado é a explicação das significantes deformações nas amplitudes das freqüências na vizinhança de uma freqüência para outra no estimador de periodograma. Este fenômeno é chamado leakage. A solução para este problema é o chamado filtragem dos dados.

3.4.6 - Filtragem dos Dados

O objetivo do filtragem dos dados é modificar a equação (3.46), que expressa a relação entre a densidade espectral estimada $P(f_k)$ a uma freqüência discreta e a densidade espectral real obtida da função contínua $P(f)$ nas proximidades dos picos de freqüência.

Quando se seleciona N pontos amostrados de uma densidade espectral estimado, na realidade está efetuando-se a multiplicação dos dados amostrados c_j por uma função janela no domínio do tempo, que vale sempre zero exceto durante o tempo total de amostragem $N\Delta$, onde assume o valor unitário. Em outras palavras, os dados estão filtrados por uma função retangular. Pelo teorema da convolução, equação (3.30), intercambiando f por t , a FFT do produto dos dados com a função janela retangular é igual a convolução da FFT dos dados com a FFT da função janela. De fato a equação (3.46) nada mais é do que a FFT da função janela retangular unitária.

$$W(s) = \frac{1}{N^2} \left[\frac{\sin(\pi s)}{\sin(\pi s / N)} \right]^2 = \frac{1}{N^2} \left| \sum_{j=0}^{N-1} e^{2\pi i s k / N} \right|^2 \quad (3.47)$$

A razão para o significativo leakage, para grandes valores de s , é que a função janela retangular altera bruscamente o seu valor; a transformada de Fourier, nesse caso, necessita de um número grande de termos de altas freqüência para representar o sinal. Para contornar tal situação, multiplica-se os dados amostrados $b(t) = \{b_0, b_1, \dots, b_{N-1}\}$ por uma função janela w_j , de forma a suavizar a transição do zero para o máximo e do máximo para o zero tal que $j = 0, 1, \dots, N$. Neste caso, as equações do estimador periodograma torna-se:

$$D_k = \sum_{j=0}^{N-1} b_j w_j e^{2\pi i j k / N} \quad k = 0, 1, \dots, N-1 \quad (3.48)$$

$$P(0) = P(f_0) = \frac{1}{W_{ss}} |D_0|^2$$

$$P(f_k) = \frac{1}{W_{ss}} \left[|D_k|^2 + |D_{N-k}|^2 \right] \quad k = 1, 2, \dots, \frac{N}{2} - 1 \quad (3.49)$$

$$P(f_c) = P(f_{N/2}) = \frac{1}{W_{ss}} |D_{N/2}|^2$$

onde W_{ss} é o somatório da janela retangular,

$$W_{ss} \equiv N \sum_{j=0}^{N-1} w_j^2 \quad (3.50)$$

f_k é dado pela equação (3.45).

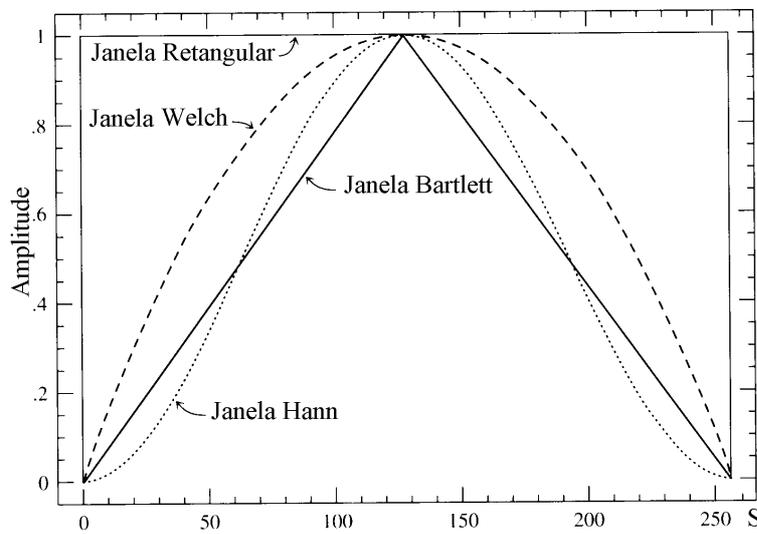


Figura (3.3) - Funções janela mais usadas na estimação da densidade de potência espectral via FFT

A forma mais geral da equação (3.46) pode ser escrita em função dos termos da função janela w_j ;

$$W(s) = \frac{1}{W_{ss}} \left| \sum_{k=0}^{N-1} w_k e^{2\pi i s k / N} \right|^2 \approx \frac{1}{W_{ss}} \left| \int_{-N/2}^{N/2} \cos(2\pi s k / N) w(k - N/2) dk \right|^2 \quad (3.51)$$

onde o termo aproximado é o estimador de uso prático, aplicável a qualquer tipo de função janela, e para $s \ll N$ (casos de interesse prático: leakage aceitável).

Algumas das funções janela mais comuns são mostradas pela figura (3.3).

$$w_i = 1 - \left| \frac{j - N/2}{N/2} \right| \equiv \text{janela Bartlett} \quad (3.52)$$

$$w_i = \frac{1}{2} \left[1 - \cos\left(\frac{2\pi j}{N}\right) \right] \equiv \text{janela Hanning} \quad (3.53)$$

$$w_i = 1 - \left(\frac{j - N/2}{N/2} \right)^2 \equiv \text{janela Welch} \quad (3.54)$$

A figura (3.4) mostra a amplitude do leakage para os tipos de janela mencionados.

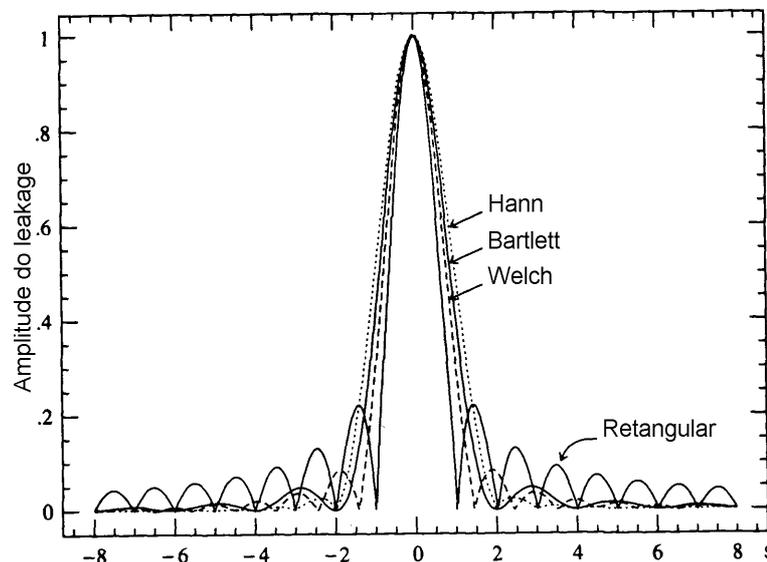


Figura (3.4) - Leakage das funções janela.

A frequência do sinal centrado no offset zero, afeta a sua vizinhança, com relação a sua amplitude.

3.5 - Transformada de Wavelet

A transformada de wavelet de um sinal $x(t)$ é um exemplo de decomposição, dilatação e translação ao longo do eixo do tempo, de uma função de análise dita wavelet, nos domínios do tempo e freqüência. A transformada de wavelet de uma função contínua é definida pela equação (3.55).

$$W_s(a,b) = \frac{1}{\sqrt{a}} \int_{-\infty}^{+\infty} x(t) g^* \left(\frac{t-b}{a} \right) dt \quad (3.55)$$

onde:

a parâmetro de dilatação que define a amplitude da janela de análise.

b parâmetro de localização da função wavelet no domínio do tempo.

g^* é o complexo conjugado da função wavelet básica.

Considerando a função wavelet (Farge, 1992) conhecida como wavelet de Morlet definida pela equação (3.56)

$$g(t) = e^{j\omega_0 t} e^{-\frac{t^2}{2}} \quad (3.56)$$

onde ω_0 é a freqüência da função wavelet.

Alternativamente a formulação da transformada contínua de wavelet pode ser obtida convertendo a função do sinal $x(t)$ e a função wavelet $g(t)$ para o domínio da freqüência,

$$W_g(a,b) = \sqrt{a} \int_{-\infty}^{+\infty} X(\omega) G^*(a\omega) e^{j\omega b} d\omega \quad (3.57)$$

onde $X(\omega)$ e $G^*(a\omega) e^{j\omega b}$ são a transformada de Fourier das funções $x(t)$ e $g^*((t-b)/a)$ respectivamente. Esta formulação da transformada de wavelet pode ser expressada discretamente como mostra a equação (3.58), facilitando sua implementação.

$$W(m,n) = \sqrt{m\Delta a} \sum_n X(f_n) G^*(m\Delta a f_n) e^{j2\pi f_n n \Delta b} \quad (3.58)$$

onde f_n é a freqüência discreta e, Δa e Δb são parâmetros de incrementos discretos respectivamente de dilatação e de translação.

Considere a função wavelet de Morlet dilatada no domínio da freqüência:

$$G(a\omega) = e^{-(a\omega - \omega_0)^2} \quad (3.59)$$

Esta função possui valor máximo quando $a\omega = \omega_0$, ou seja

$$a = \frac{\omega_0}{2\pi f} \quad (3.60)$$

assim a transformada de wavelet fornece uma representação do sinal através de uma simples transformação linear entre os domínios do tempo e frequência, tal como a transformada inversa de Fourier.

3.5.1 - Módulo e Fase da Transformada de Wavelet

A transformada de wavelet discreta de um sinal resulta numa matriz complexa cujo módulo e fase pode ser facilmente representado. Considere o sinal harmônico dado pela equação (3.61)

$$x(t) = X_0 \cos(\omega_0 t) \quad (3.62)$$

cujo espectro é dado pela equação (3.63), onde δ é a função delta de Dirac.

$$X(\omega) = \frac{X_0}{2\pi} (\delta(\omega - \omega_n) + \delta(\omega + \omega_n)) \quad (3.63)$$

O espectro da função wavelet de Morelet dilatada e transladada é definida como:

$$G_{a,b}(\omega) = e^{-(a\omega - \omega_0)^2} e^{j\omega b} \quad (3.64)$$

Usando a propriedade da função delta de Dirac

$$\int_{-\infty}^{+\infty} \delta(t - t_0) f(t) dt = f(t_0) \quad (3.65)$$

e a formulação da transformada de wavelet no domínio da frequência dada pela equação (3.55), a transformada de wavelet da equação (3.62), pode ser escrito como:

$$W(a,b) = \sqrt{a} X_0 e^{-(a\omega_n - \omega_0)^2} e^{j\omega_n b} \quad (3.66)$$

Fixando o valor da dilatação ($a = a_i$), o módulo e a fase da transformada de wavelet para um valor fixo de frequência são:

$$\begin{aligned} |W(a_i, b)| &= \sqrt{a_i} X_0 e^{-(a_i \omega_n - \omega_0)^2} \\ \langle W(a_i, b) &= \omega_n b \end{aligned} \quad (3.67)$$

A equação (3.67) ilustra como o módulo é diretamente proporcional à amplitude do sinal

harmônico, enquanto que a fase é dada pela variação da fase do sinal multiplicada pelo tempo. Considerando o sinal mais geral:

$$x(t) = k(t) \cos(\varphi(t) t) \quad (3.68)$$

onde $k(t)$ é a função envelope e $\varphi(t)$ é a função de fase. Neste caso a transformada de wavelet de $x(t)$ pode ser expressado como:

$$W(a,b) = \sqrt{a} k(t) e^{-(a\varphi(t)-\omega_0)^2} e^{j\varphi(t)b} \quad (3.69)$$

Fazendo ($a = a_i$) tem-se que:

$$\begin{aligned} |W(a,b)| &= \sqrt{a_i} X_0 e^{-(a_i\varphi(t)-\omega_0)^2} \\ \langle W(a,b) &= \varphi(t) b \end{aligned} \quad (3.70)$$

A equação anterior mostra como uma função $x(t)$, com uma função envelope geral $k(t)$ ou função fase $\varphi(t)$, dependente do tempo, pode ser determinada usando o modulo e a fase da transformada de wavelet, para um valor fixo de freqüência.

3.5.2 - Aplicação: Estimação da freqüência Natural e Razão de Amortecimento.

A resposta impulsiva de um sistema mecânico amortecido, de um grau de liberdade, é dado pela equação (3.71),

$$x(t) = A e^{-\zeta\omega_n t} \cos(\omega_d t + \phi_0) \quad (3.71)$$

onde A é a magnitude residual, ω_n é a freqüência não amortecida, ω_d é a freqüência amortecida ($\omega_d = \omega_n \sqrt{1-\zeta^2}$), ϕ_0 é a fase e ζ é a razão de amortecimento. O sinal $x(t)$ também pode ser representado como um vetor no plano complexo:

$$z(t) = A e^{-\zeta\omega_n t + j(\omega_d t + \phi_0)} \quad (3.72)$$

ou seja, $x(t) = Re[z(t)]$.

Uma das técnicas de estimação de parâmetros modais (Felman, 1994) é baseada na aproximação do sinal complexo $z(t)$ com o correspondente sinal analítico $z_a(t)$ definido como:

$$z_a(t) = x(t) + jH[x(t)] \quad (3.73)$$

onde $j = \sqrt{-1}$ e $H[x(t)]$ é a transformada de Hilbert de $x(t)$. Esta aproximação é exata se $z(t) = z_a(t)$. A transformada de Hilbert permite estimar as funções envelope e fase, que para um

sistema amortecido, de um grau de liberdade, são dadas por:

$$\begin{aligned} |z_a(t)| &\equiv Ae^{-\zeta\omega_n t} \\ \langle z_a(t) &\equiv \omega_d t + \phi_0 \end{aligned} \quad (3.74)$$

Define-se frequência instantânea de um sinal como:

$$\varphi(t) = \frac{1}{2\pi} \frac{d}{dt} (\langle z_a(t) \rangle) \quad (3.75)$$

neste caso a inclinação da linha da função fase representa a frequência angular amortecida

$$\varphi(t) \equiv \omega_d \quad (3.76)$$

e a inclinação da função envelope é definido como razão de decaimento σ .

$$\sigma = \zeta\omega_n \quad (3.77)$$

Se usarmos transformada de wavelet ao invés da transformada de Hilbert, considerando o sinal dado pela equação (3.71), a partir das equações (3.69) e (3.70), tem-se:

$$\begin{aligned} k(t) &= \frac{|W(a_i, b)|}{\sqrt{ae^{-(a_i\varphi(t)-\omega_0)^2}}} = Ae^{-\zeta\omega_n t} \\ \varphi(t) &= \langle W(a_i, b) = \omega_d t + \phi_0 \end{aligned} \quad (3.78)$$

Segue-se então o procedimento análogo quando utilizado a transformação de Hilbert pode ser usado para a determinação dos parâmetros do sistema.

Capítulo 4

Fundamentos de Redes Neurais

4.1- Introdução

Quase a totalidade da literatura referente a dinâmica de estrutura contém como premissa básica a linearização do seu modelo. Indicativos disto são os termos "frequências naturais" e "modos de vibrar" que são propriedades de sistema lineares.

Na prática, todos os sistemas vibratórios contém quase sempre não-linearidades, principalmente provenientes de amortecimentos, o que pode distanciar, de forma comprometedora, os resultados experimentais e teóricos do sistema linear "equivalente".

Pesquisadores de dinâmica não linear (Cawley, 1979), tem conseguido progressos na modelagem de sistemas; porém com um alto grau de complexidade, o que não convence os projetistas a usarem tal formulação.

Entretanto, em certas aplicações tais como interação fluido-estrutura e dinâmica de rotação, a influência de uma fonte excitadora as vezes só pode ser corretamente percebida e quantificada quando se considera as não linearidades presentes no sistema. Neste contexto, a técnica de Redes Neurais é para os dinamicistas algo como impossível de não experimentar. Ela mapeia qualquer não linearidade e não requer modelagem formal do sistema a ser tratado.

O aparecimento das redes neurais no meio científico ocorreu na década de 40 (McCulloch e Pitts, 1943) quando foi apresentado um modelo simplificado de neurônio biológico, com vistas à aplicação em circuitos eletrônicos. No final da década de 70, foram identificadas deficiências no modelo proposto (Minsky e Papert, 1969a). Este fato levou muitos pesquisadores, ávidos por soluções imediatas para os seus problemas, a abandonarem o desenvolvimento desta ferramenta.

O grande impulso no estudo das redes neurais aconteceu a partir da década de 80, graças ao desenvolvimento de computadores de maior capacidade de processamento e a importantes descobertas teóricas tais como o "Error Back-propagation"; essa combinação tem revelado as redes neurais como ferramentas altamente promissoras e que poderão revolucionar o modo com que a ciência atualmente aborda vários problemas.

Redes neurais artificiais (Liu e Mengel, 1992) foram usadas para monitorar, diagnosticar e classificar defeitos de juntas esféricas usadas em máquinas ferramentas e robôs. Neste trabalho o valor RMS do sinal, o valor da máxima amplitude do sinal no domínio da frequência e o espectro de potência foram usados para compor os dados para alimentar a rede neural. Seis estados distintos da junta esférica foram considerados. A rede neural treinada foi capaz de distinguir uma junta esférica defeituosa de uma não defeituosa em 100% de acerto dos casos. A classificação

quanto aos seis defeitos considerados foi obtida com 97% de sucesso. Aspectos importantes como a estrutura da rede e o tamanho do conjunto de dados usados na fase de treinamento foram discutidos.

Redes neurais artificiais (Liu e at all, 1995) foram usadas para gerar uma tabela de valores discretos de frequências naturais de um dado sistema, a partir de medições de sua função resposta em frequência. Foram usadas duas redes neurais, a primeira descobre o número de ressonâncias dentro de uma dada banda de frequência e a segunda determina o valor pico de cada ressonância. As duas redes foram montadas usando o algoritmo de Back-propagation. As estratégias de treinamento testadas foram Windrow (Windrow, 1985) e Levenberg-Marquardt (Vanderplaats, 1984). Para aplicações de um e dois graus de liberdade. Os resultados foram considerados altamente satisfatórios.

Redes neurais unidirecionais (Neto, 1997) foram usadas para modelar sistemas dinâmicos não lineares. Os resultados da aplicação de filtro de Kalman foram usados no processo de treinamento da rede, uma vez que este pode ser considerado como uma estimação de parâmetros estocásticos, em substituição ao tradicional algoritmo dos mínimos quadrados.

Redes neural recorrentes (Green e Manchada, 1996) foram investigadas quanto a sua capacidade de identificar sistemas dinâmicos não lineares. Limitando as redes neurais recorrentes para um ordem fixa, apenas algumas classes de redes foram capazes de aproximar a identificação do sistema.

Uma metodologia para automação do processo de detecção de defeitos em máquinas rotativas usando sinais de vibração foi proposta (Uhl, 1994). Os sinais de vibração das partes móveis e expostas a desgaste e falhas, são compactados por uma rede neural de base radial e posteriormente analisados por uma rede neural back-propagation, que é responsável pela classificação do tipo de falha. São analisados os resultados de um soprador industrial. Os teste (Uhl, 1994) apresentaram bons resultados e a metodologia mostrou-se promissora.

Redes neurais (Mistry e Nair, 1994) foram empregadas para identificação e controle de sistemas não lineares cujo comportamento é parcialmente conhecido. O controle via redes neurais mostrou-se eficiente e requer menos informação do sistema que as teorias de controles convencionais.

Redes neurais (Ball e at all, 1992) foram usadas para localizar e quantificar uma falha numa estrutura metálica. Como treinamento da rede foram usados dados de deformação de diferentes pontos da estrutura, gerados por modelo de elementos finitos, representativo de um sistema real. Tal técnica mostrou-se eficiente.

Redes neurais com o algoritmo de back-propagation (Qi e at all, 1994) foram empregadas na identificação de parâmetros de modelo dinâmico de estruturas.

A literatura relatou (Lucifredi e Hou, 1996) alguns problemas de construção de redes neurais para identificação de sistemas lineares e bilineares, e em especial atenção para identificação de razão de rigidez de sistemas bilineares, que exibem características típicas de

falhas em sistemas dinâmicos. Foi treinada uma rede para identificar não só a magnitude da razão de rigidez mas também especificar sua posição. São mostrados aplicações para sistemas de um, dois e três graus de liberdade. Os resultados mostraram promissores o uso de redes neurais para detecção e localização de falhas estruturais.

4.2- O que são e quais as aplicações das redes neurais?

As redes neurais são modelos computacionais ou sistemas de processamento de informações que, de um modo geral, podem ser tratados como um dispositivo "caixa preta" que aceita uma entrada e produz uma saída.

Existem três situações básicas nas quais o emprego das redes neurais é vantajoso:

- Situações onde se requer poucas decisões e que envolvam grandes quantidades de dados (processamento de imagens, etc.).
- Situações onde as não-linearidades são fundamentais e predominantes no modelo (controle de robôs, etc.).
- Situações onde soluções muito próximas do ótimo devem ser fornecidas rapidamente (sistema de aero-navegação automática, etc.).

As redes neurais, potencialmente, podem ser utilizadas para as mais diversas finalidades dentre as quais destacam-se:

- **classificação**: a rede neural fornece como saída uma classe representativa de um dado padrão fornecido como entrada.
- **reconhecimento de padrões**: se um determinado padrão é fornecido à rede, esta será capaz de reproduzi-lo como saída.
- **complementação de padrões**: um padrão incompleto é fornecido à rede neural e esta reproduzirá o mesmo padrão completando as partes faltantes.
- **remoção de ruído**: um padrão contaminado é apresentado à rede neural e ela será capaz de reproduzir o sinal removendo a maior parte, senão a totalidade, do ruído.
- **otimização**: um conjunto de condições iniciais de um específico problema de otimização é apresentado à rede neural que, produzirá como saída um conjunto de condições que representa a solução ótima para o problema.
- **controle**: um padrão representativo do estado atual do sistema e da resposta desejada, é fornecido à rede neural que produzirá como saída uma seqüência de comandos que possibilitará a obtenção da resposta desejada.

4.3- Conceitos básicos

Uma rede neural compõe-se, basicamente, dos seguintes entidades:

- um conjunto de unidades de processamento, também chamadas de células ou neurônios, agrupados em camadas;
- uma regra de propagação que determinará como que as informações provenientes da camada anterior, se comporá numa entrada efetiva i para os neurônios;
- uma função de ativação F_i , que produzirá o nível de ativação a_i do neurônio, a partir da entrada efetiva i ,
- uma entrada externa ou um offset θ para cada neurônio e
- um método de armazenamento de informações.

4.3.1- Unidade de processamento: neurônio

O neurônio é elemento responsável pelo processamento numérico da rede. Suas características básicas, para a maioria das aplicações, são:

- atua de forma isolada dos demais neurônios,
- realiza apenas operações matemáticas predefinidas,
- só tem acesso às informações locais e
- produz somente um valor de saída

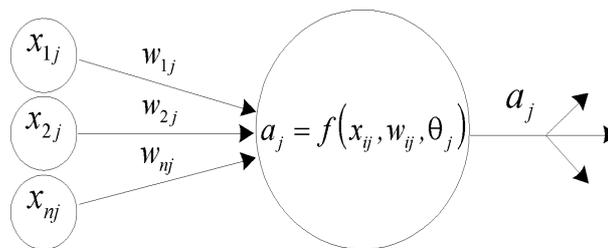


Figura (4.1) - Modelo gráfico de um neurônio

4.3.2- Ligações entre neurônios: sinapses.

São caminhos lógicos por onde os sinais fluem de um neurônio para outro. Cada sinapse, em particular, tem um peso característico (w_{ij}) que amplifica ou atenua o sinal transportado (x_{ij}), como mostra a figura (4.1).

Existem algumas aplicações (Minsky e Papert, 1969b) que utiliza sinapses duplas entre os neurônios. Uma das ligações transportam a variância dos sinais de entrada do neurônio,

enquanto a outra ligação transporta um sinal representativo da média dos pesos das sinapses que chegaram ao neurônio da camada i , como mostra a figura (4.2). Este tipo de ligação é conhecido como média-variância.

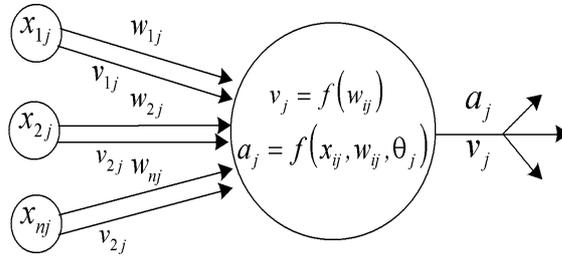


Figura (4.2) - Modelo gráfico de uma sinapse dupla

4.3.3- Função de ativação

As funções de ativação converte um domínio infinito para um outro domínio finito, delimitado por uma faixa de valores predeterminados. A maioria das redes neurais empregam cinco tipos de funções de ativação, representadas na figura (4.3):

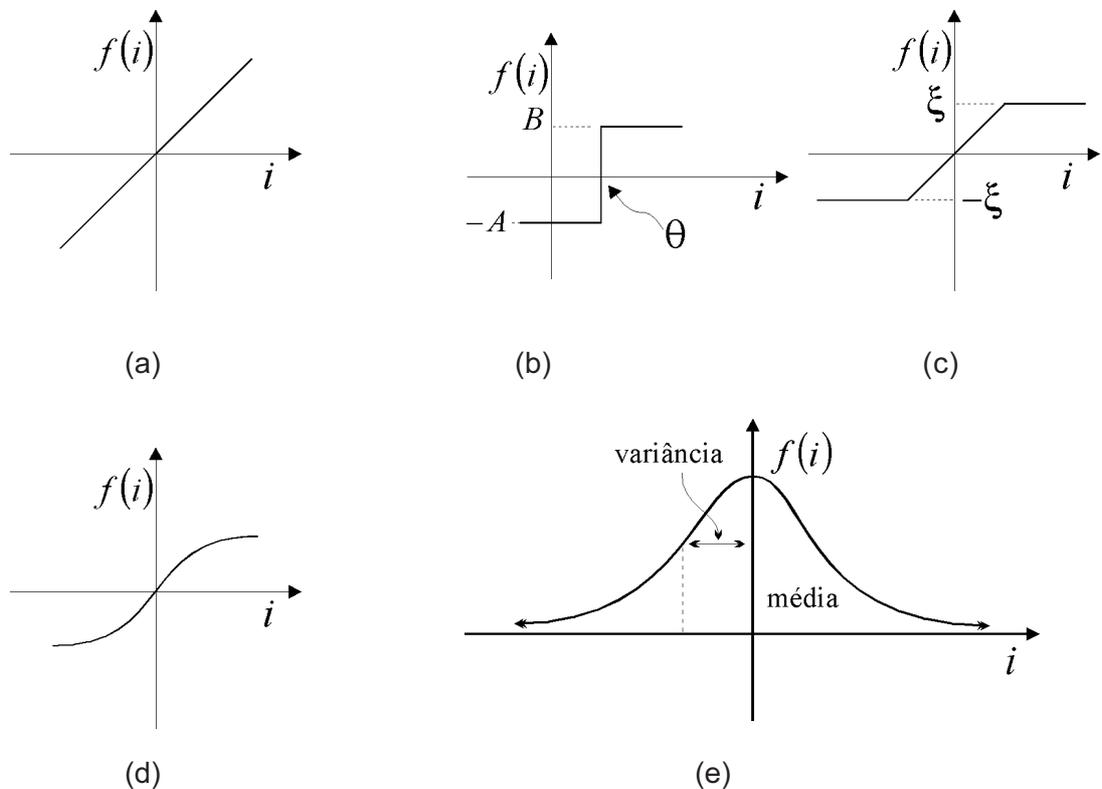


Figura (4.3) - Funções de ativação. (a) Função linear. (b) Função degrau. (c) Função rampa. (d) Função sigmoide. (e) Função gaussiana

Função linear

A função de ativação linear, como mostra a figura (4.3.a), produz uma saída linear, a partir de uma entrada i , descrita pela equação seguinte.

$$a = f(i) = ki \quad (4.1)$$

onde k é um escalar positivo.

Função degrau

A função de ativação tipo degrau, como mostra a figura (4.3.b), produz somente dois valores, A e B , escalares positivos. Se o valor do sinal de entrada i é igual o maior que um dado valor de referência θ , então a função de ativação produz o valor B , por outro lado produz o valor $-A$. Matematicamente esta função pode ser descrita como:

$$a = f(i) = \begin{cases} B & \text{se } i \geq \theta \\ -A & \text{se } i < \theta \end{cases} \quad (4.2)$$

É comum nas redes neurais do tipo Hopfield (Hopfield, J. J., 1982a), usar função degrau bipolar, ou seja,

$$a = f(i) = \begin{cases} 1 & \text{se } i \geq \theta \\ -1 & \text{demais casos} \end{cases} \quad (4.3)$$

Função Rampa

A função de ativação tipo rampa, como mostra a figura (4.3.c), é uma combinação das funções degrau e linear. A função rampa produz valores que variam linearmente entre dois pontos de saturação simetricamente dispostos em torno da origem; fora dessa faixa, produz valores constantes.

$$a = f(i) = \begin{cases} \xi & \text{se } i \geq \xi \\ i & \text{se } |i| < \xi \\ -\xi & \text{se } i \leq -\xi \end{cases} \quad (4.4)$$

onde ξ é o valor de saturação da função. Os pontos $i = \xi$ e $i = -\xi$, são descontinuidades na função f .

Função sigmoidal

A função de ativação sigmoidal, como mostra a figura (4.3.d), é uma versão contínua da função rampa. A função sigmoide é uma função monotônica, não decrescente, tem um domínio finito e produz respostas não-lineares dentro de uma faixa predefinida. A função sigmoidal mais comum é matematicamente definida como:

$$a = f(i) = \frac{1}{1 + e^{-\xi i}} \quad (4.5)$$

onde $\xi > 0$ (geralmente $\xi = 1$), promove um valor da saída na faixa ente 0 e 1. Ainda, dois outros tipos de funções alternativas a função sigmoidal são encontrados, a função tangente hiperbólica, que produz valores na faixa de +1 a -1,

$$a = f(i) = \tanh(i) \quad (4.6)$$

e a função das razões quadráticas aumentada, que produz valores na faixa entre 0 e 1.

$$a = f(i) = \begin{cases} \frac{i^2}{1+i^2} & \text{se } i > 0 \\ 0 & \text{demais casos} \end{cases} \quad (4.7)$$

Função gaussiana

A função de ativação gaussiana, vista na figura (4.3.e), é simétrica em relação a origem e requer um valor predefinido de variância $\nu > 0$.

A maioria das aplicações que utilizam esta função de ativação, o faz em conjunto com as ligações, ou sinapses do tipo dual, ou média-variância.

$$a = f(\bar{i}) = e^{\left(\frac{-(\bar{i})^2}{\nu}\right)} \quad (4.8)$$

onde \bar{i} é o valor médio da entrada e ν é a variância previamente definida.

4.4- Fundamentos

Os fundamentos das redes neurais consistem na compreensão dos seus conceitos matemáticos básicos, usados para descrever e analisar o seu processamento, os seus mecanismos de funcionamento e entender sua terminologia.

O estudo de uma rede neural pode ser feito compreendendo seus três princípios elementares:

A topologia: a organização da rede neural;

O aprendizado: como os pesos são atualizados;

A recuperação de informações: como a rede recupera uma informação previamente armazenada.

4.4.1- Topologia da Rede Neural

A topologia das redes neurais, refere-se a maneira com a qual a rede neural está organizada: o número de camadas, o tipo de ligação entre os neurônios e como os dados fluem através dela.

Quanto a forma de conexão, dois tipos básicos se distingue:

Rede neural unidirecional (feed-forward networks)

Aqui o fluxo de dados é estritamente unidirecional, ou seja os dados entram na rede pela camada de entrada, percorrem as camadas invisíveis, onde são processados e saem pela camada de saída da rede. Nenhum neurônio recebe sinal de outro da mesma camada, e tão pouco de uma camada subsequente.

Rede neural recorrente (feedback networks)

Ao contrário da rede unidirecional, os neurônios são retroalimentados e os valores das funções de ativação sofrem um processo de relaxação.

Rede Neural Multicamadas Unidirecional

Uma rede unidirecional obrigatoriamente é estruturada em camadas. Neste caso cada camada recebe os sinais provenientes da camada anterior, processa-os e fornece sinais de saída para a camada subsequente, estabelecendo assim o fluxo de informações.

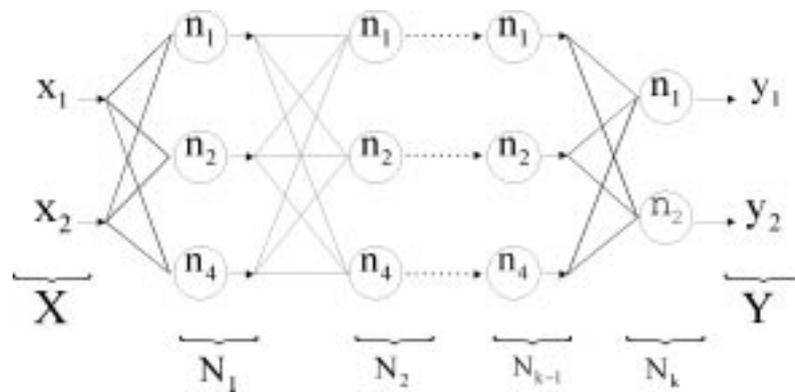


Figura (4.4) - Rede unidirecional multicamadas.

Não existe ligações entre neurônios de uma mesma camada. A figura (4.4) mostra graficamente a estrutura desta rede.

4.4.2- Treinamento das Redes Neurais

A principal característica desta ferramenta numérica, que é a rede neural, é a sua capacidade de aprender. Neste contexto, a aprendizagem é definida como a variação dos peso das ligações, de modo a armazenar implicitamente uma dada informação, que poderá ser facilmente recuperada.

Os procedimentos de aprendizagem podem ser separados em duas situações distintas:

- **O aprendizado supervisionado ou associativo:** neste caso, a rede é treinada, pela ação de um supervisor que fornece a ela, pares de padrões de entrada e de resposta, representativos de um dado fenômeno.

- **O aprendizado não supervisionado ou auto-organizado:** aqui, a rede é treinada para responder a padrões de características diferentes. Não se tem acesso aos sinais de saída do sistema para formar o par entrada - saída. A rede neural terá que perceber as sutilezas presentes em cada um dos padrões, agrupando-os em categorias distintas. Este método de aprendizagem será visto no estudo das redes neurais auto-organizadas.

A Regra Delta

Uma generalização do algoritmo perceptron de treinamento foi apresentado em 1960 (Windrow e Hoff, 1960) o algoritmo de aprendizagem "least mean square" (LMS) também conhecido como regra delta, tem sido aplicado para redes com uma única camada com funções de ativação puramente lineares.

Seja a saída de um neurônio dada por:

$$a = \sum_j w_j x_j + \theta \quad (4.9)$$

A função erro é obtida pela média do somatório dos erros quadráticos:

$$E = \sum_p E^p = \frac{1}{2} \sum_p (d^p - a^p)^2 \quad (4.10)$$

Onde:

p é um padrão de entrada,

E^p representa o erro para o padrão p ,

d^p é a saída desejável para o padrão p ,

a^p é a saída obtida para o padrão p .

Os valores de todos os pesos de ligações são encontrados minimizando a função erro através do método do gradiente descendente. A idéia é fazer variações nos pesos, proporcional ao negativo da derivada do erro medido em relação a cada peso, para o corrente padrão de entrada.

$$\Delta w_j \Big|_p = -\gamma \frac{\partial E^p}{\partial w_j} \quad (4.11)$$

onde γ é uma constante de proporcionalidade. A derivada é dada por:

$$\frac{\partial E^p}{\partial w_j} = \frac{\partial E^p}{\partial a^p} \frac{\partial a^p}{\partial w_j} \quad (4.12)$$

Como a equação (4.9) é linear, tem-se que:

$$\frac{\partial a^p}{\partial w_j} = x_j \quad (4.13)$$

Logo:

$$\frac{\partial E^p}{\partial a^p} = -(d^p - a^p) \quad (4.14)$$

De modo que:

$$\Delta w_j \Big|_p = \gamma \delta^p x_j \quad (4.15)$$

onde $\delta^p = (d^p - a^p)$ é a diferença entre a saída desejada e a saída obtida para o padrão p .

O Algoritmo Back-propagation

Minsky e S. Papert (Minsky e S. Papert, 1969) mostraram que as redes unidirecional com mais de duas camadas podem superar muitas restrições, mas não apresentaram a solução para o problema de como ajustar os pesos das sinapses a partir da camada de entrada para as camadas invisíveis. A resposta para esta questão foi apresentada por Rumelhart (Rumelhart, Hinton e Williams 1986).

A idéia central desse algoritmo é que o erro de cada camada invisível é determinado retrocedendo os erros a partir da camada de saída. Por essa razão este algoritmo é denominado back-propagation learning rule.

O algoritmo back-propagation pode ser considerado uma generalização da regra delta para funções de ativação não-lineares e rede neural multicamadas.

A Regra Delta Generalizada

Aqui a Regra Delta será generalizada com relação ao número de camadas da rede e à função de ativação dos neurônios que abrangerá funções não-lineares.

Seja uma função de ativação diferenciável com relação ao sinal de entrada total, dada por:

$$a_i^p = F(i_i^p) \quad (4.16)$$

onde

$$i_j^p = \sum_j w_{ij} a_j^p + \theta_i \quad (4.17)$$

A generalização da correção dos pesos através da regra delta apresentado anteriormente, torna-se:

$$\Delta w_{ij}|_p = -\gamma \frac{\partial E^p}{\partial w_{ij}} \quad (4.18)$$

O erro E^p é definido como o erro quadrático total para o padrão p medido na camada de saída da rede N_0 .

$$E^p = \frac{1}{2} \sum_j^{N_0} (d_j^p - a_j^p)^2 \quad (4.19)$$

A medida do erro global é dada por:

$$E = \sum_p E^p \quad (4.20)$$

Pela regra da cadeia pode-se escrever que:

$$\frac{\partial E^p}{\partial w_{ij}} = \frac{\partial E^p}{\partial i_i^p} \frac{\partial i_i^p}{\partial w_{ij}} \quad (4.21)$$

Da equação (4.17) tem-se que:

$$\frac{\partial i_i^p}{\partial w_{ij}} = a_i^p \quad (4.22)$$

Definido

$$a_i^p = \frac{\partial E^p}{\partial i_i^p} \quad (4.23)$$

obtem-se uma função de atualização dos pesos, resultado de um gradiente descendente numa superfície:

$$\Delta w_{ij} \Big|_p = \gamma \delta_i^p a_j^p \quad (4.24)$$

O termo δ_i^p pode ser obtido para cada camada i da rede de forma recursiva através da propagação dos sinais de erro.

Para calcular o termo δ_i^p , pode-se usar a regra da cadeia, assim δ_i^p será expresso pelo produto de dois fatores, o primeiro reflete a variação do erro em função do sinal de saída e o segundo fator reflete a variação do sinal de saída em função da variação do sinal de entrada da camada em questão:

$$\delta_i^p = -\frac{\partial E^p}{\partial i_i^p} = -\frac{\partial E^p}{\partial a_i^p} \frac{\partial a_i^p}{\partial i_i^p} \quad (4.25)$$

Da equação (4.16) tem-se que:

$$\frac{\partial a_i^p}{\partial i_i^p} = F_i'(i_i^p) \quad (4.26)$$

que é simplesmente a derivada da função de ativação F da i ésima camada, calculada para o conjunto de sinais de entrada i_i^p . Para calcular o primeiro termo da equação (4.25), considera-se dois casos. O primeiro, assume que a camada i é a camada de saída da rede N_o , neste caso

segue-se a definição de E^p , que é dado por:

$$\frac{\partial E^p}{\partial a_i^p} = -(d_i^p - a_i^p) \quad (4.27)$$

Substituindo este resultado na equação (4.26) tem-se que

$$\delta_i^p = (d_i^p - a_i^p) F'_i(i_i^p) \quad (4.28)$$

para qualquer neurônio da camada i . O segundo caso é quando a camada i não é a camada N_o da rede neural. Aqui não se conhece diretamente a contribuição do erro dessa camada na composição do erro total da rede. Entretanto o erro medido pode ser escrito como uma função dos sinais de entrada para as várias camadas invisíveis,

$$E^p = E^p(i_1^p, i_2^p, i_3^p, \dots, i_i^p) \quad (4.29)$$

Usando a regra da cadeia, pode-se escrever:

$$\frac{\partial E^p}{\partial a_i^p} = \sum_{h=1}^{N_o} \frac{\partial E^p}{\partial i_i^p} \frac{\partial i_i^p}{\partial a_i^p} = \sum_{h=1}^{N_o} \frac{\partial E^p}{\partial i_i^p} \frac{\partial \left(\sum_{k=1}^{N_k} w_{hk} a_k^p \right)}{\partial a_i^p} = \sum_{h=1}^{N_o} \frac{\partial E^p}{\partial i_i^p} w_{hi} = - \sum_{h=1}^{N_o} \delta_h^p w_{hi} \quad (4.30)$$

Substituindo esta equação na equação (4.25), tem-se que:

$$\delta_i^p = F'_i(i_i^p) \sum_{h=1}^{N_o} \delta_h^p w_{hi} \quad (4.31)$$

As equações (4.28) e (4.31) possibilita um procedimento recursivo para obtenção dos termos δ para todas as camadas da rede neural, que são usados para calcular a variação dos pesos de ligação através da equação (4.24). Este procedimento constitui a regra delta generalizada para uma rede neural multicamadas unidirecional, de funções de ativação não-lineares.

4.5- Trabalhando com Back-propagation

A aplicação da regra delta generalizada consiste de duas fases:

Na primeira fase os sinais de entrada i_i^p são apresentados e propagados unidirecionalmente através da rede neural para calcular os valores de saída a_i^p dos neurônios de cada camada. Esta saída é comparada com o valor da saída desejada, resultando num sinal de erro δ_i^p para cada neurônio das várias camadas da rede.

Na segunda fase a rede neural é percorrida no sentido inverso, onde o sinal do erro é

passado para todos os neurônios de cada camada e a variação nos pesos de ligação é apropriadamente calculado.

4.5.1- Ajuste dos pesos aplicado à função de ativação sigmoidal

O peso da ligação entre o neurônio i de uma camada e o neurônio da camada subsequente, é ajustado por um valor proporcional ao produto do sinal de erro δ pelo valor do sinal de saída enviado ao longo da ligação.

$$\Delta w_{ij}|_p = \gamma \delta_i^p a_j^p \quad (4.32)$$

Se o neurônio em questão pertence a uma camada de saída N_0 da rede neural, o sinal de erro é dado por:

$$\delta_i^p = (d_i^p - a_i^p) F'_i(i_i^p) \quad (4.33)$$

A função de ativação F sigmoidal é definida como:

$$a_i^p = F(i_i^p) = \frac{1}{1 + e^{-i_i^p}} \quad (4.34)$$

Neste caso a derivada da função é dada por

$$F'(i_i^p) = \frac{\partial}{\partial i_i^p} \left(\frac{1}{1 + e^{-i_i^p}} \right) = \frac{1}{(1 + e^{-i_i^p})} (-e^{-i_i^p}) = \frac{1}{(1 + e^{-i_i^p})} \frac{e^{-i_i^p}}{(1 + e^{-i_i^p})} = a_i^p (1 - a_i^p) \quad (4.35)$$

de modo que o sinal de erro para um neurônio da camada N_0 pode ser escrito como

$$\delta_i^p = (d_i^p - a_i^p) a_i^p (1 - a_i^p) \quad (4.36)$$

O sinal de erro para um neurônio pertencente a uma camada invisível é determinado recursivamente em termos dos sinais de erros dos neurônios com os quais tem ligações, bem como dos pesos das ligações

Para a função sigmoidal tem-se:

$$\delta_i^p = F'(i_i^p) \sum_{h=1}^{N_o} \delta_h^p w_{hi} = a_i^p (1 - a_i^p) \sum_{h=1}^{N_o} \delta_h^p w_{hi} \quad (4.37)$$

4.5.2- Taxa de aprendizado e o termo de inércia

O procedimento de aprendizado requer que a variação dos pesos seja proporcional a $E^p / \partial w$. O verdadeiro método do gradiente descendente requer passos infinitesimais. A constante de proporcionalidade é a taxa de aprendizagem λ , que na prática é arbitrado e o seu valor é escolhido o maior possível desde que não permita oscilações na aprendizagem. Uma maneira de evitar oscilações para valores grandes de λ , é vincular a variação dos pesos em relação a seus calores anteriores, adicionando um termo de inércia:

$$\Delta w_{ij}(t+1) = \lambda \delta_i^p a_j^p + \alpha \Delta w_{ij}(t) \quad (4.38)$$

onde:

t simboliza o número da atualização,

α determina o efeito da atualização passada.

A função do termo de inércia é mostrado na figura (4.5). Quando o termo de inércia não é usado, um tempo maior é necessário para que o ponto de mínimo seja atingido usando uma baixa taxa de aprendizado; entretanto se for usada uma alta taxa da aprendizado, o ponto de mínimo nunca será alcançado devido às oscilações. Adicionando o termo de inércia o mínimo é localizado rapidamente.

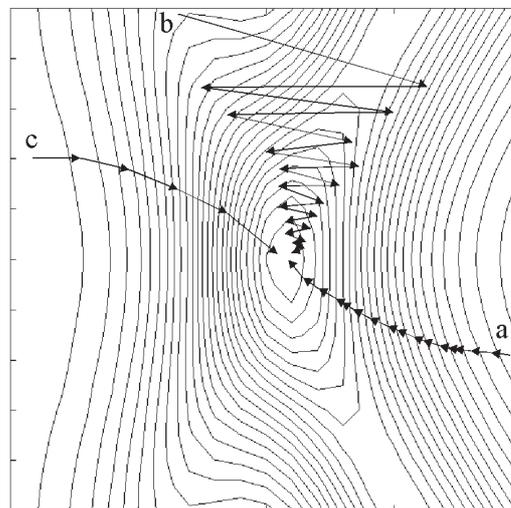


Figura (4.5) - O ajuste dos pesos. (a) - pequenas taxas de aprendizagem. (b) grandes taxas de aprendizagem. (c) - grandes taxas de aprendizagem e termo de inércia adicional.

4.5.3 - Deficiências do método de Back-propagation.

Apesar de que o algoritmo de aprendizagem de Back-propagation esteja fundado num método matemático sólido, de minimização de erros (método do gradiente descendente), existem alguns aspectos que não garantem a sua aplicabilidade universal. O mais incomodo é o longo processo de treinamento. Isto pode ser resultado por valores inadequados de η , a taxa de aprendizagem, e de δ , o fator de inércia. A maioria dos métodos avançados de treinamento, baseado no método de Back-propagation, têm algum dispositivo de ajuste automático da taxa de aprendizagem. Outras falhas de treinamento têm origem em duas fontes conhecidas: os mínimos locais e a chamada paralisia neural.

Paralisia neural

Em consequência do processo de treinamento da rede, os pesos podem ser ajustados para valores muito grandes; por essa razão o valor i (o sinal total de entrada), bem como o sinal de saída a , de um neurônio pertencente a uma camada invisível, pode assumir valores, em módulo, de grande magnitude, o que torna o sistema fortemente instável e o processo de aprendizagem não entra em rota de convergência. Isto pode se exemplificado para o caso em que a função de ativação seja a sigmoideal, onde os pesos são atualizados por um fator proporcional a $a_i^p (1 - a_i^p)$ que seguramente é próximo de zero. A este fenômeno chama-se paralisia neural.

Mínimo local

A função erro de um rede neural é uma superfície complexa repleta de picos e vales; isto leva o método de minimização de erros a se tornar refém de um mínimo local, não conseguindo atingir o mínimo global da função erro. Alguns métodos probabilísticos podem diminuir ou remover este obstáculo, mas o processo de minimização tenderá ser mais lento. Uma outra solução possível é aumentar o número de camadas invisíveis; isto elevará significativamente o número de dimensões da superfície da função erro, como consequência, o incremento do número de mínimos locais é de tal sorte que, para um dado valor de passo (η), o método do gradiente descendente se torna robusto o suficiente para sair dos vales menores.

Aprendizagem viciada.

Embora, teoricamente, o algoritmo de Back-propagation realiza o método do gradiente descendente sobre o erro total de um único padrão, pertencente a um conjunto de padrões representativos de um dado evento ou fenômeno, e não simultaneamente para o um conjunto de padrões ($p = 1, 2, \dots, P$), existem indicações empíricas que mostram que o erro global para o conjunto converge rapidamente.

Existem situações em que a seqüência dos padrões, usados durante o processo de treinamento, faz com que a rede neural fique "viciada" ou seja, ela reconhece preferencialmente os primeiros padrões a ela apresentados. Este comportamento é tipicamente encontrado quando se deseja treinar uma rede para reconhecer um dado fenômeno, o qual sofre forte variação durante o período em que se adquiriu o seu conjunto de padrões característico. Este tipo de problema pode ser eliminado apresentando à rede, durante o treinamento, os padrões fora da seqüência em que foram adquiridos; outra solução proposta por vários autores e que, ao se efetuar o processo de treinamento da rede usar algoritmos diferentes intercaladamente entre um padrão e outro.

Ajuste inadequado

Redes neurais, cujos algoritmos de aprendizagem utilizam métodos de minimização de erros, são altamente sensíveis (Rumelhart, et al, 1986) ao número de neurônios que compõem as camadas invisíveis. Em tais métodos, o aceite ou não do resultado do aprendizado é determinado por uma medida global dos erros gerados por cada uma das experiências integrantes do banco de dados. Portanto, tais métodos não garantem que cada uma das experiências foram isoladamente bem ajustadas.

Uma rede neural com poucos neurônios nas camadas invisíveis, devidamente aceita como 'treinada', quase sempre apresenta o fenômeno do subajuste; por outro lado, se o número de neurônios das camadas invisíveis for excessivo, o fenômeno que se manifesta é o sobreajuste. Seja o caso de três redes neurais, multicamadas, back-propagation, cujas arquiteturas são 5x5x1, 5x10x1 e 5x40x1, treinadas para um mesmo banco de dados. A figuras (4.6), (4.7) e (4.8), mostram o resultado da execução das três redes.

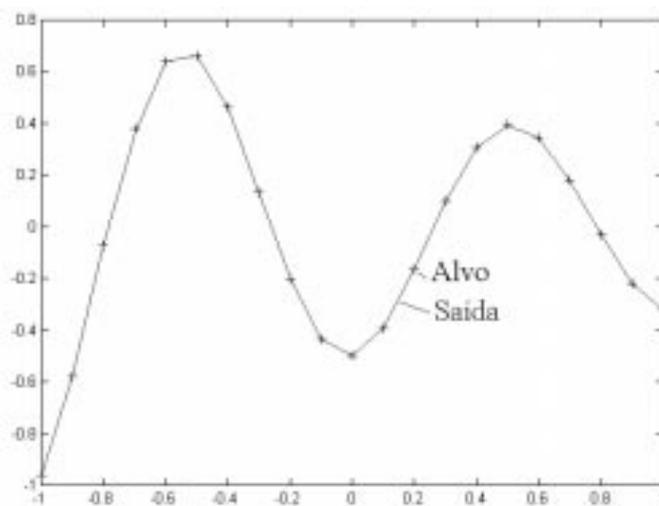


Figura (4.6) - Arquitetura apropriada. 5x10x1 neurônios.

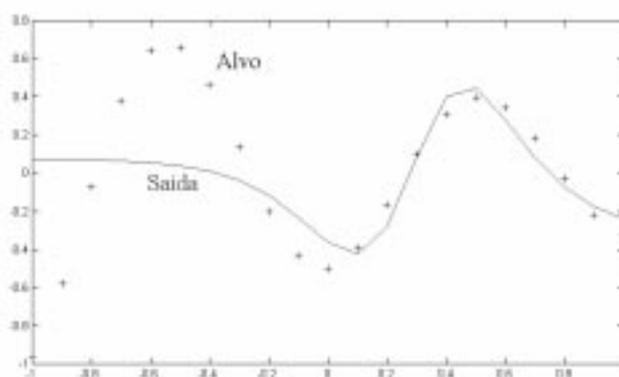


Figura (4.7) - Arquitetura 5x5x1 neurônios: fenômeno do subajuste.

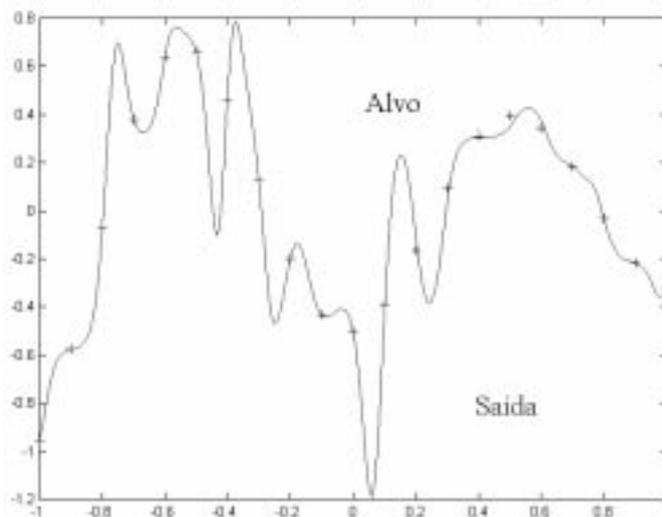


Figura (4.8) - Arquitetura 5x40x1 neurônios: fenômeno do sobreajuste.

Não existe um receituário capaz de indicar a arquitetura apropriada para um dado caso. O número de neurônios bem como a função de transferência, dependem da natureza matemática de cada caso.

4.6 - Algoritmos avançados

Atualmente, muitos pesquisadores estão empenhados em incrementar o mesmo em expandir as características básicas do algoritmo de Back-propagation. Algumas novas modificações neste método serão aqui discutidos, porém ainda não se têm condições de avaliar se tais mudanças serão de fundamental importância ou se estão fadadas ao fracasso.

A evolução mais óbvia é trocar o método do gradiente descendente, que possui passo fixo, pelo método das direções conjugadas (Powell, M. J. D. 1984). Diferentemente do método do gradiente descendente que minimiza uma função f sempre numa mesma direção u , neste método, em cada passo do processo de minimização, uma nova direção perpendicular a u é estabelecida.

Suponha a seguinte aproximação por série de Taylor, da função a ser minimizada:

$$f(x) = f(p) + \sum_i \left. \frac{\partial f}{\partial x_i} \right|_p x_i + \frac{1}{2} \sum_i \left. \frac{\partial^2 f}{\partial x_i \partial x_j} \right|_p x_i x_j + \dots \cong \frac{1}{2} x^T A x - b^T x + c \quad (4.39)$$

Onde

\mathbf{T} denota transposição,

$$\begin{aligned} [A]_{ij} &\equiv \left. \frac{\partial^2 f}{\partial x_i \partial x_j} \right|_p \\ b &\equiv -\nabla f|_p \\ c &\equiv f(p) \end{aligned} \quad (4.40)$$

$[A]$, uma matriz simétrica positiva definida, é a matriz Hessiana de f e p . O gradiente de f é dado por:

$$\nabla f = Ax - b \quad (4.41)$$

de modo que uma variação em x resulta na alteração do gradiente tal que:

$$\delta(\nabla f) = A(\delta x) \quad (4.42)$$

Assumindo que f foi minimizada na direção de u_i até um ponto onde o gradiente $-g_{i+1}$ de f tornou-se perpendicular a u_i , ou seja,

$$u_i^T g_{i+1} = 0 \quad (4.43)$$

e é buscada então, uma nova direção g_{i+1} . Isto é feito movendo-se ao longo de u_{i+1} . Para que a minimização efetuada na direção u_i não seja danificada, requer que o novo gradiente da função f também seja perpendicular a direção u_i ,

$$u_i^T g_{i+2} = 0 \quad (4.44)$$

Combinando as equações (4.43) e (4.44) tem-se que:

$$0 = u_i^T (g_{i+1} - g_{i+2}) = u_i^T \delta(\nabla f) = u_i^T A u_{i+1} \quad (4.45)$$

Os vetores u_i e u_{i+1} , que satisfazer a equação (4.45), são ditos conjugados.

O processo de minimização aqui, consiste nos seguintes passos:

- Partindo de um ponto p_0 , na direção u_0 , tal que o gradiente seja $g_0 = -\nabla f(p_0)$, resulta num novo ponto p_1 . Para $i \geq 1$, o cálculo da nova direção é dado por

$$u_{i+1} = g_{i+1} + \gamma_i u_i \quad (4.46)$$

onde γ_i é escolhido de modo que as sucessivos gradientes sejam perpendiculares, ou seja,

$$\gamma_i = \frac{g_{i+1}^T g_{i+1}}{g_i^T g_{i+1}} \quad (4.47)$$

com $g_k = -\nabla f|_{p_k}$ para qualquer $k \geq 0$.

- O passo seguinte é o cálculo de $p_{i+2} = p_{i+1} + \lambda_{i+1} u_{i+1}$ onde u_{i+1} é escolhido de tal modo a minimizar $f(p_{i+2})^3$.

Este processo acima descrito é conhecido como Método de Fletcher-Reeves (Vanderplaats, G. 1984) (Powell, M. J. D., 1977).

Outras modificações, no método de Back-propagation, baseadas numa taxa de aprendizagem adaptativa e independente para cada peso, tem aprestado resultados promissores (Silva, F. M. e Almeida, L. B., 1990)

4.7 - Redes Neurais Auto-organizadas

Os demais tipos de redes neurais, para serem treinadas, necessitam pares de dados tipo causa-efeito. Entretanto, existem situações onde esse par de informação não é disponível, tem-se acesso apenas ao conjunto de padrões x_p , relativo as causas. Neste caso o volume de informações é importante e o treinamento é feito levando em conta as características redundantes

das amostras.

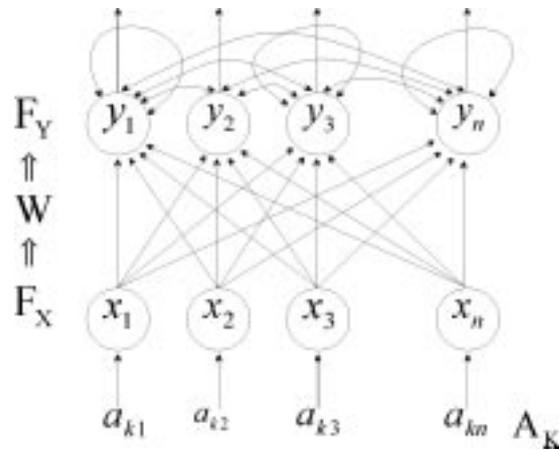


Figura (4.9) - Redes auto-organizadas

Aqui, o processo de treinamento é realizado sem a presença de um supervisor e o processo de atualização dos pesos é geralmente baseado numa forma de competição entre os neurônios.

Alguns exemplos típicos, dessa classe de problema, são:

- **Classificação:** a rede neural, atuando sobre um conjunto de padrões, deve classificar e agrupar os padrões semelhantes em classes distintas. Após o treinamento, dado um padrão, a rede fornece o rótulo da sua classe característica.
- **Quantização vetorial:** esse problema ocorre quando uma discretização ótima deve ser feito sobre um espaço contínuo.

4.7.1 - Classificação de padrões

A aprendizagem competitiva

Este tipo de aprendizagem, introduzida por Grossberg (Grossberg, S., 1982), é um método criado para classificar automaticamente, um conjunto de padrões. A aprendizagem competitiva é um procedimento de dois passos que acopla o processo de atualização dos pesos de ligação e o processo de recuperação das informações armazenadas em uma rede neural de duas camadas.

Cada neurônio da camada de entrada recebe um componente do padrão de entrada, e cada neurônio da camada de saída da rede, fornece uma classe diferente.

Passo 1: determinação do neurônio vitorioso ou a classe vitoriosa. Um padrão é transmitido, unidirecionalmente, através das conexões entre os neurônios da camada A_k , para os neurônios da camada F_Y , utilizando a seguinte regra:

$$y_j = \sum_{i=1}^n x_i w_{ij} \quad (4.48)$$

onde:

x_i são os neurônios da camada F_X , $i=1,2,\dots,n$

y_j são os neurônios da camada F_Y , $j=1,2,\dots,n$

w_{ij} é a matriz dos pesos entre x_i e y_j .

Seja $W_j = \{w_{1j}, w_{2j}, \dots, w_{nj}\}$ o conjunto dos pesos que chegam até a camada de saída da rede. Então W_j é um vetor de referência e representa a classe j .

Com a finalidade de evitar inconvenientes numéricos (overflow e underflow), tanto o padrão de entrada A_k ($k=1,2,\dots,m$) quanto o vetor W_j são normalizados fazendo a norma euclidiana igual a unidade, o que levará à seguinte relação:

$$0 \leq \left(y_j = A_k W_j = \sum_{i=1}^n a_{ki} w_{ij} \right) \leq 1 \quad (4.49)$$

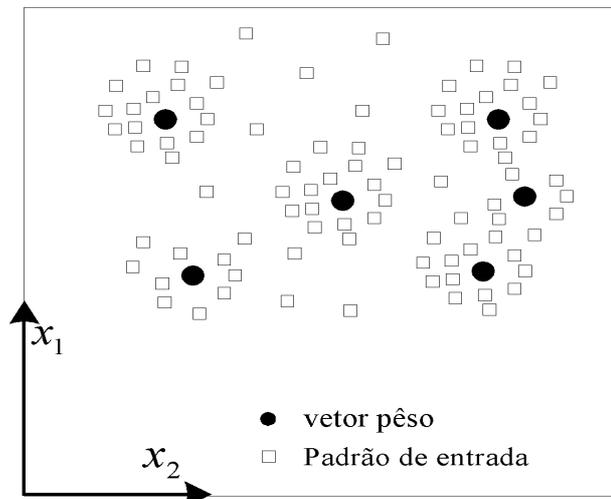


Figura (4.10) - Classes e centróides

Os valores de y_j são usados como valores iniciais para o processo iterativo conhecido como o vencedor tudo. O resultado desse processo é a identificação do neurônio vencedor, ou seja, o neurônio da camada de saída que produz o maior valor de ativação. Isto é feito usando a seguinte equação:

$$y_j = \begin{cases} 1 & \text{se } y_j > y_k \text{ para todos } j \neq k \\ 0 & \text{nos demais casos} \end{cases} \quad (4.50)$$

Passo 2: Ajuste dos pesos das conexões do neurônio vencedor. Neste tipo de

aprendizagem, somente o vetor de referência, ou seja, os pesos que chegam até o neurônio vencedor, é ajustado segundo a uma regra de atualização:

$$w_{ij}(t+1) = \frac{w_{ij}(t) + \rho(t) \cdot y_j(a_{ki}(t) - w_{ij}(t))}{\|w_{ij}(t) + \rho(t) \cdot y_j(a_{ki}(t) - w_{ij}(t))\|} \quad (4.51)$$

onde $\rho(t)$ é uma função do tempo, decendente não nula. O vetor W_j , após o termino do processo de ajuste, representa o centróide da classe j . Este tipo de aprendizado é principalmente empregada nas rede neurais auto-organizadas.

Quantização vetorial

Uma outra aplicação da aprendizagem competitiva é a quantização vetorial. O objetivo aqui é dividir o espaço de entrada em subespaços disjuntos, estabelecendo para cada vetor de entrada o rótulo do subespaço ao qual pertence, como mostrado na figura (4.8).

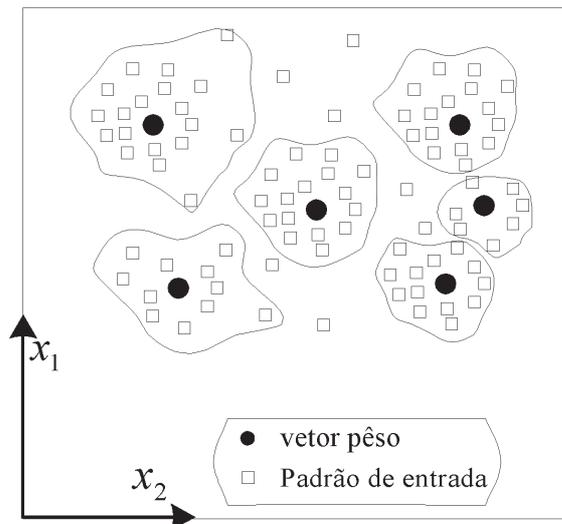


Figura (4.11) - Subespaços disjuntos

Ao contrário do processo de agrupamento, não se está interessado em encontrar padrões similares mas sim determinar o espaço que um conjunto particular de padrões similares exerce seu domínio.

Para este caso, o processo de treinamento pode ser sumarizado da seguinte forma:

Passo 1 - Um padrão x é apresentado à rede. Os valores de ativação são calculados $a_{ij} = F(x, w)$ para uma matriz de pesos aleatoriamente estabelecida.

Passo 2 - Para cada vetor de pesos $w_i = (w_1, w_2, \dots, w_n)$ calcula-se a distância euclidiana do padrão de entrada x para encontrar o neurônio vencedor k .

Passo 3 - A atualização dos pesos é feita pela equação a seguir:

$$w_{ij}(t+1) = \frac{w_{ij}(t) + \rho(t) \cdot y_j(a_{ki}(t) - w_{ij}(t))}{\|w_{ij}(t) + \rho(t) \cdot y_j(a_{ki}(t) - w_{ij}(t))\|} \quad (4.52)$$

Passo 4 - Um passo de treinamento supervisionado é realizado.

$$w_{hk}(t+1) = w_{hk}(t) + \gamma \left(\sum_i a_i w_{hi} - w_{hk}(t) \right) \quad (4.53)$$

Onde:

k refere-se ao neurônio vencedor,
 γ é taxa de aprendizagem e
 $\sum_i a_i w_{hi}$ é saída desejável.

Se uma função $g(x, k)$ é definida tal que

$$g(x, k) = \begin{cases} 1 & \text{se } k \text{ for vencedor} \\ 0 & \text{demais casos} \end{cases} \quad (4.54)$$

é possível mostrar que o processo converge para

$$w_{hi} = \int_{\mathbb{R}^n} \sum_i (a_i w_{hi}) g(x, k) dx. \quad (4.55)$$

4.8 - Redes Recorrentes

Nas redes anteriormente discutidas, todos os dados fluem numa única direção, após um número de passos finito e conhecido, todos os valores de ativação (a_{ij}) da rede são conhecidos. Os pesos da rede converge para um estado desejável por meio de uma regra de aprendizagem.

Nas chamadas redes recorrentes os valores de ativação sofrem um processo de relaxação, porem os pesos permanecem inalterados. Neste tipo de rede existem conexões retroalimentadoras que ligam os neurônios da camada de saída aos da camada de entrada da rede. Quando o estado da rede ($[a_{ij}]$ e $[w_{ij}]$) é instável, os neurônios entram em funcionamento até convergirem para um estado estável.

Hopfield (Hopfield, J. J., 1982b) reuniu as idéias referentes a este tipo de rede e apresentou uma completa análise matemática, por isso este tipo de rede é geralmente chamada de rede de Hopfield.

4.8.1 - A Rede de Hopfield

A rede de Hopfield, como mostra a figura (4.9), consiste de n neurônios, interconectados, dispostos de forma anelar, que atualizam seus valores de ativação assíncrono e independentemente dos outros neurônios. Todos os neurônios são, simultaneamente, de entrada e de saída. Os valores de ativação são binários. Originalmente Hopfield escolheu valores de ativação de 0 a 1, mas valores -1 e 1 apresentam vantagens numéricas.

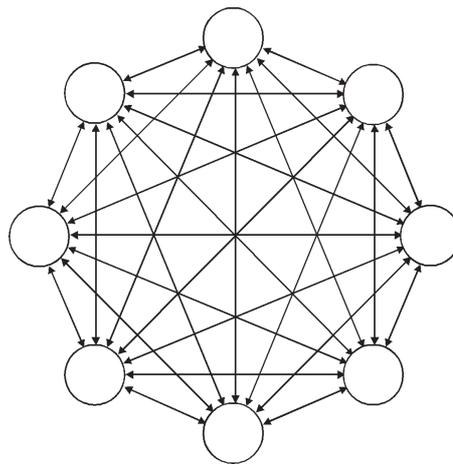


Figura (4.12) - Rede neural anelar

O estado da rede é dados pelos valores de ativação $[a_{ij}]$. A entrada para o neurônio i , $a_i(t+1)$ ciclos é dado pela soma dos pesos.

$$i_i(t+1) = \sum_{j \neq i} a_j(t) w_{ij} + \theta_i \quad (4.56)$$

Um filtro aplicado na entrada do neurônios para obter uma nova função de ativação $a_i(t+1)$ no ciclo $(t+1)$.

$$a_i(t+1) = \begin{cases} +1 & \text{se } i_i(t+1) > U_i \\ -1 & \text{se } i_i(t+1) < U_i \\ a_i(t) & \text{demais casos.} \end{cases} \quad (4.57)$$

Por uma questão de simplicidade escolhe-se $U_i = 0$, mas não é uma imposição.

Um neurônio de Hopfield é dito estável quando para um dado ciclo t em concordância com as equações (4.56) e (4.57),

$$\text{senal}(a_i(t)) = \text{senal}(i_i(t-1)) \quad (4.58)$$

Um estado α é dito estável se todos os neurônios estão estáveis.

Para garantir a estabilidade da rede nos pontos de sela, Hopfield (Hopfield, J. J., 1982c) mostrou que é necessário remover as ligações bidirecionais $w_{ij} = w_{ji}$, então o comportamento o sistema pode ser descrito pela função energia

$$\varepsilon = -\frac{1}{2} \sum_{i \neq j} a_i a_j w_{ij} - \sum_i \theta_i a_i \quad (4.59)$$

4.9 - Redução Dimensional: Método Estatístico

Para a maioria das aplicações envolvendo redes neurais, o número de variáveis potencialmente importantes é gigantesco.

Existem dois grandes problemas quando se processa uma aplicação com um número elevado de variáveis. O primeiro, é o custo computacional, o que implica na necessidade de máquinas com grande capacidade de memória e de processamento. O segundo problema é quando existe uma grande quantidade de variáveis presentes na amostra e poucas amostras compõem o banco de dados disponível para o treinamento da rede.

Existe ainda um outro problema potencialmente importante associado a um grande número de variáveis, a existência de uma substancial correlação entre as variáveis da amostra, ou seja, quando parte significativa das variáveis são interdependentes.

É tanto possível ter total dependência entre duas variáveis, quanto é possível independência para um centena delas; mas grande é a probabilidade de existir interdependência significativa entre as variáveis para casos com elevado número de variáveis.

A formação do banco de dados disponibilizado para o treinamento das redes neurais, influencia fortemente o funcionamento, seja pelo excesso ou pela fraca presença de informações significativas. Em ambos os casos o custo computacional é penalizado e o treinamento pode ser inviabilizado.

Será discutido aqui, um método que possibilita extrair as informações proveitosas presentes num conjunto numeroso de variáveis, deixando de lado as informações redundantes. O resultado esperado é um conjunto de menor número de variáveis e com baixo nível de interdependência entre elas.

Como ponto de partida, o embasamento estatístico (não mostrado aqui) dessa metodologia garante que a maioria das informações importantes são preservadas. Existe sempre a possibilidade de uma informação preciosa ser inadvertidamente deixada de lado. Entretanto, felizmente esses casos são pouco frequentes (Master, Timothy, 1995).

4.9.1 - Componentes Principais (cp)

Os componentes principais são números gerados a partir de uma coleção de vários experimentos, de modo de encontrar a maior variância de cada experimento, com o menor número possível de variáveis.

Cada cp é um escalar calculado como o produto escalar entre uma amostra e um vetor de pesos previamente definido.

Seja \mathbf{v} o vetor peso previamente definido tal que $\mathbf{v} = (v_1, v_2, v_3, \dots, v_m)$. Seja $\mathbf{x} = (x_1, x_2, x_3, \dots, x_m)$ uma simples amostra. Define-se o cp tal que:

$$y = \mathbf{v}^T \mathbf{x} = v_1 x_1 + v_2 x_2 + v_3 x_3 + \dots + v_m x_m \quad (4.60)$$

Tal equação consiste de um único componente principal, uma variável escalar.

Seja m o número de variáveis mensuradas. Na prática sempre se trabalha com p componentes principais de interesse tal que $1 \leq p < m$. A equação (4.60) é então modificada.

$$\mathbf{y} = \mathbf{V}^T \mathbf{x} \quad (4.61)$$

Onde $[\mathbf{y}]_{m \times p}$ é uma matriz que define p componentes principais.

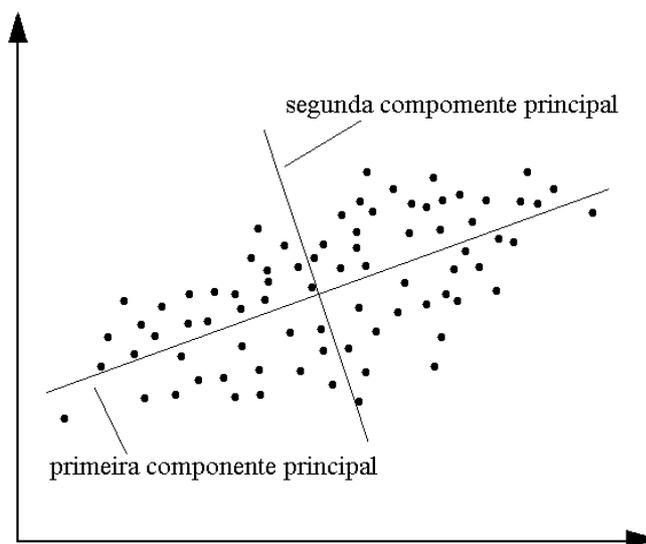


Figura (4.13) - Dois componentes principais

Suponha que queiramos extrair um único cp de uma amostra de dados. Em termos matemáticos, nos queremos definir \mathbf{v} de modo que a variância de $\mathbf{v}^T \mathbf{x}$ seja maximizado com relação ao universo de vetores possíveis. A figura (4.13) mostra a interpretação geométrica da condição de máxima variância.

Seja um gráfico xy com linha representativas das centróides. As projeções dos pontos

sobre a linha centróide representam o produto $\mathbf{v}^T \mathbf{x}$.

Seja v_1 a definição do vetor componente principal. Define-se v_2 como sendo o vetor componente secundário. O vetor v_2 deve expressar a máxima variância através da coleção de dados, bem como deve eliminar redundância das variáveis mensuradas, este fato implica que o segundo componente principal seja independente do primeiro. Em termos geométricos, v_2 deve ser ortogonal a v_1 .

Este procedimento pode ser repetido tantas vezes quanto se fizer necessário. Cada cp sucessivo é calculado de tal modo que ele contenha a máxima variância possível, sujeito a condição de não correlação com todos os prévios componentes principais calculados.

Conhecida as propriedades dos vetores componentes principais, faz-se necessário estabelecer um algoritmo para calcular a matriz V , cujas p colunas serão multiplicadas com as linhas de dados (Anderson, T. W. 1958). Como todo procedimento estatístico, o primeiro passo é calcular a média. Seja x_{ij} o valor medido para a variável número j , amostra número i da coleção de n amostras, logo

$$\mu_j = \frac{1}{n} \sum_{i=1}^n x_{ij} \quad (4.62)$$

Para um conjunto de treinamento, a matriz de covariância $[W]_{m,m}$ é simétrica e sua diagonal principal contém as variâncias de cada uma das m variáveis mensuradas. Isto pode ser representado pela equação (4.63).

$$C_{ij} = C_{ji} = \frac{1}{n} \sum_{k=1}^n (x_{ki} - \mu_i)(x_{kj} - \mu_j) \quad (4.63)$$

Os vetores que definem os cp são os autovetores da matriz de covariância, e seus autovalores são as variâncias. A operação consiste então em calcular os autovalores e autovetores da matriz C .

O autovetor correspondente ao autovalor de maior módulo, ocupa a primeira coluna de V , e assim sucessivamente. A adição de colunas na matriz V é interrompido tão logo que o módulo do autovalor correspondente tornar-se pequeno. Isto é, a sua contribuição para a variação do conjunto de treinamento, torna-se inexpressiva.

Como se determina p , o número de cp (número de colunas de V)? Não existe uma resposta única para esta questão. A soma dos autovalores de C é igual a soma das variâncias das variáveis originais (diagonal de C). Por essa razão é perfeitamente razoável considerar uma fração da variância total como sendo representativa do conjunto original.

Uma boa prática é expressar graficamente a fração da variância total (de 0 até 1) como função de p . À medida que se adiciona cp ao novo conjunto de treinamento, a soma dos seus

autovalores é incrementada. Por inspeção visual, quando a variação desse incremento não for significativa, para-se o processo. Eis então, o número p . A figura (4.14) ilustra este procedimento.

Define-se razão da variância como sendo a relação entre o somatório dos autovalores retidos e o somatório dos autovalores totais da matriz covariância.

$$\text{Razão da variância} = \frac{\sum \lambda^2_{\text{retido}}}{\sum \lambda^2_{\text{total}}} \quad (4.64)$$

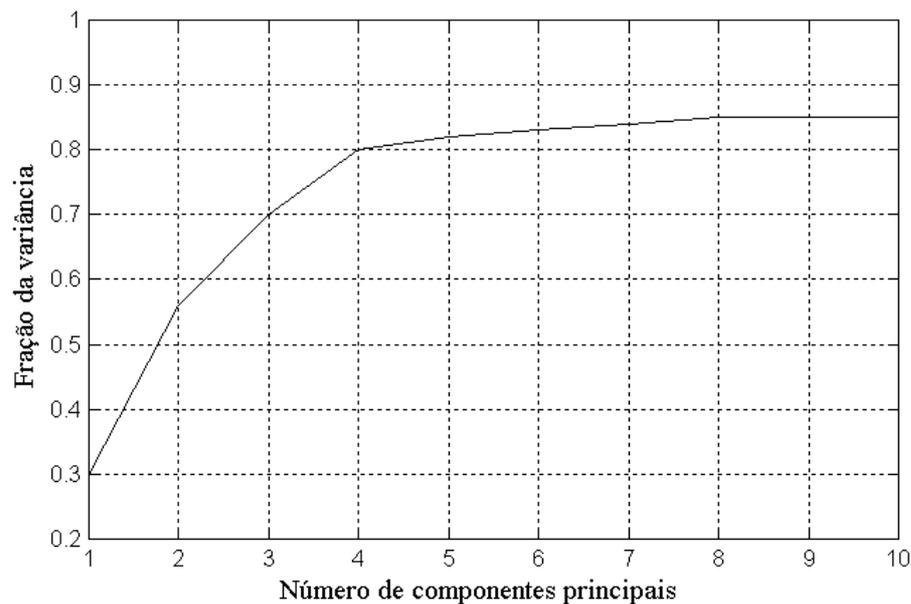


Figura (4.14) - Influência do número de componentes principais

De um modo geral é suficiente extrair 90 por cento da variância original, desprezando assim os demais componentes principais. Porém, se alguma informação importante estiver fracamente representada no sinal mensurado, certamente ela fará parte dos componentes principais não consideradas, pois seus autovalores são pequeno módulo. Nestes casos, fatalmente ocorrerão problemas no processo de treinamento da rede neural. Só a experiência e sensibilidade do profissional pode identificar e solucionar tais transtornos. Felizmente esses casos são raros.

4.9.2 - Escalonamento e alterações numéricas

Foi visto que os componentes principais são obtidos calculando os autovalores e autovetores da matriz de covariância de um conjunto de dados, usando a equação (4.63). Obtém-se então a matriz V (matriz peso) retendo parte dos autovetores que possuam autovalores significativos. Então através da equação (4.61) obtém-se y , um novo conjunto de dados, ditos compactados.

Algumas alterações nessa técnica podem trazer benefícios numéricos significativos. A primeira alteração, diz respeito à centralização dos componentes principais. O uso direto da equação (4.61) fornece variáveis cuja média é geralmente diferente de zero. Para aplicações como redes neurais, é interessante que seus valores de entrada sejam centralizados. Neste caso, os componentes principais são calculados pela equação (4.65)

$$\mathbf{y} = \mathbf{V}^T(\mathbf{x} - \boldsymbol{\mu}) \quad (4.65)$$

onde $\boldsymbol{\mu} = (\mu_1, \mu_2, \dots)$.

A próxima alteração envolve o escalonamento dos componentes principais.

Em determinadas aplicações de redes neurais, cujos dados de treinamento que enfatizam a importância de alguns poucos primeiros componentes, pode acarretar baixa taxa de aprendizado e o aumento da probabilidade dos métodos de minimização de erros, ficarem aprisionados em algum um mínimo local.

Existem situações onde, mais de uma componente principal tem variância de mesmo módulo. O escalonamento para a variância unitária é então uma boa prática. Isto é feito dividindo cada componente principal calculado pela raiz quadrada de seu autovalor correspondente, como mostra a equação (4.66).

$$\mathbf{y} = \frac{1}{\sqrt{L}} \mathbf{V}(\mathbf{x} - \boldsymbol{\mu}) \quad (4.66)$$

Em determinadas situações, onde existam duas ou mais variáveis de valores numéricos muito distintos (por exemplo, comprimento: 1.2 metros e peso: 10000 gramas), pertencentes a uma mesma amostra, certamente levará a autovalores de módulo igualmente distintos. Assim, o autovalor de módulo menos expressivo, pode ser inadvertidamente desprezado. Nesses casos, pode-se obter C escalonado, usando variáveis que não foram previamente escalonadas, usando a equação (4.67).

$$C'_{ij} = \frac{C_{ij}}{\sqrt{C_{ii}C_{jj}}} \quad (4.67)$$

4.9.3 - Componentes principais de um grupo de centróides

A metodologia anteriormente discutida, tem como propósito a extração da máxima variação de um conjunto de dados, expressando-a com um número menor de variáveis independentes. Mas, esses casos são bastante particulares, quando examinados à luz do universo de possibilidades.

Seja o caso em que uma amostra é formada de grupos de classes de variáveis distintas, e ainda, o objetivo da rede neural seja o de classificar tais variáveis nas suas diversas classes. Aqui é falho a aplicação da metodologia previamente discutida.

Nos problemas de classificação existem dois grupos de casos a serem considerados. O primeiro caso é quando existe a presença de ruído, uma fonte de variação que não se relaciona a nenhuma das classes presentes no sinal. É simplesmente uma variação natural que ocorre dentro de cada classe. Este caso é chamado de variância intraclasse. O segundo caso, chamado de variância interclasse, é quando existe verdadeiramente diferenças entre as classes, mas suas variáveis tendem a aglutinar-se em áreas próximas ou na mesma área.

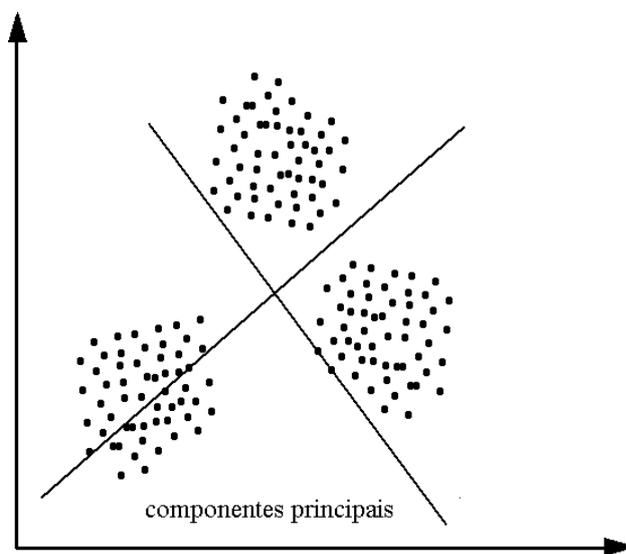


Figura (4.15) - Formação dos componentes principais a partir de grupos de centróides

Como calcular as componentes principais nos casos em que todas as classes se sobrepõe numa mesma área e ainda contém variáveis que não se enquadram em nenhuma classe (ruído)? Esta é uma situação nada ideal, mas bastante realista. Não existe nenhum bom método aplicável em todos os casos. A situação anteriormente descrita é então remediada assumindo que, cada uma das classes isoladamente, é unimodal, ou seja, cada uma das classes tendem aglutinar-se em uma única área, não se separando em dois ou mais subgrupos. A figura (4.15) mostra um exemplo com três classes.

Uma vez assumida a hipótese anterior, a variância interclasse pode ser aproximada pela centróide (valor médio) da classe. Então, os componentes principais são formados pelas médias das classes.

Quantos componentes principais pode-se calcular? Se existem duas classes, tem-se duas centróides. Dois pontos definem uma linha. Logo, tem-se um único componente principal. Se existem três classes, tem-se três centróides. Três pontos definem um plano e assim

sucessivamente. Porém, um número pequeno de componentes principais são suficientes para obter sucesso num processo de classificação. Caso contrario, provavelmente métodos estatísticos tradicionais sejam mais adequados que redes neurais, para o caso em questão.

4.9.4 - Interpretação do método de compactação de dados.

Considere p vetores experimentais $\{e\}$ de n elementos, representativos de um mesmo fenômeno. Seja $[P]$ uma matriz com p colunas, onde cada colunas é um experimento $\{e\}$ distinto, tal como mostra a equação (4.68).

$$[P] = \begin{bmatrix} e_{11} & \cdots & e_{p1} \\ e_{12} & \cdots & e_{p2} \\ e_{13} & \cdots & e_{p3} \\ \cdots & \cdots & \cdots \\ e_{1n} & \cdots & e_{pn} \end{bmatrix}_{n,p} \quad (4.68)$$

Define covariância a C_{xy} entre duas variáveis x e y como:

$$C_{xy} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x - \mu_x)(y - \mu_y) p(xy) dx dy = E[(x - \mu_x)(y - \mu_y)] \quad (4.69)$$

onde $p(xy)$ é a função densidade de probabilidade conjunta, das variáveis x e y . E é a esperança estatística.

O módulo da função covariância (Bendat e Piersol, 1986) é sempre menor ou igual produto dos desvios padrões das variáveis,

$$|C_{xy}| \leq \sigma_x \sigma_y \quad (4.70)$$

uma vez que o coeficiente de correlação entre as duas variáveis consideradas é no máximo igual a unidade.

$$\rho_{xy} = \frac{C_{xy}}{\sigma_x \sigma_y} \in [-1,1] \quad (4.71)$$

Se as variáveis x e y são estatisticamente independentes, implica dizer que são não correlacionadas; logo a covariância entre elas é nula ($C_{xy} = 0$). Por outro lado, covariância nula não implica que as variáveis x e y sejam estatisticamente independentes.

Se a função probabilidade conjunta $p(xy)$ for gaussiana, a nulidade da covariância ($C_{xy} = 0$) garante que as variáveis x e y são estatisticamente independentes, e ainda que x e y

são não correlacionadas.

No caso discreto, se a função densidade de probabilidade conjunta for gaussiana, tem-se que:

$$C_{ij} = E[(x_i(e) - \mu_i)(x_j(e) - \mu_j)] \quad (4.72)$$

onde:

e identifica a amostra,

$$C_{ii} = \sigma_i^2 \quad (4.73)$$

σ^2 é a variância.

A partir da equação (4.73) obtém-se:

$$[C_{ij}] = \begin{bmatrix} C_{11} & \cdots & C_{1n} \\ C_{21} & \cdots & C_{2n} \\ \cdots & \cdots & \cdots \\ C_{n1} & \cdots & C_{nn} \end{bmatrix} \quad (4.74)$$

onde C_{ij} expressa a correlação entre $x_i(e)$ e $x_j(e)$.

Pode-se determinar um novo espaço de representação para a matriz de covariância, no qual a matriz $[C_{ij}]$ é diagonalizada (direção principal). A transformação linear é obtida pela solução do problema de autovalor e autovetor da matriz $[C_{ij}]$. Neste espaço as variáveis são correlacionadas (e se o processo for gaussiano com n dimensões, está garantida a independência estatística entre todas as medidas obtidas em todos os experimentos).

Seja $[\lambda]_{n,n}$ e $[V]_{n,n}$ respectivamente as matrizes dos autovalor e autovetores de $[C_{ij}]$. O traço das matrizes de covariância e a matriz dos autovalores deve ser igual.

$$\sum[\lambda_{ii}] = \sum[C_{ii}] \quad (4.75)$$

Valores pequenos de λ_{ii} podem ser removidos da matriz $[\lambda_{ii}]$ sem que a equação (4.73) seja significativamente afetada. Ao retirar λ_{kk} , mantendo $\sum[\lambda_{ii}]_{\text{retidos}} = \alpha \sum[C_{ii}]$ onde α é um nível desejado de aproximação tal que $0 \leq \alpha \leq 1$. Neste caso são eliminados da matriz $[V]_{n,n}$ as colunas (autovetores) correspondentes aos λ_{kk} (autovalores) eliminados, alterando assim a base do espaços da transformação linear para $[\bar{V}_{ij}]$.

A matriz $[\bar{V}_{ij}]$ é então usada para transformar linearmente a matriz $[P]_{n,p}$ para um novo espaço vetorial, retendo $\alpha\%$ de significância dos dados inicialmente medidos, tal como mostra a equação (4.76).

$$[\bar{V}]_{r,n} \cdot [P]_{n,p} \Leftrightarrow [\bar{P}]_{r,p} \quad (4.76)$$

Da equação anterior, tem-se que $r < n$, logo $[\bar{P}]_{r,p}$ pode ser interpretada como uma versão compactada de $[P]_{n,p}$.

Esta é a metodologia proposta neste trabalho para a compactação de dados.

Capítulo 5

Bancada de Dinâmica de Rotação

5.1 - Introdução

Parte substancial deste trabalho constituiu-se do projeto, construção, montagem e caracterização de uma bancada, destinada ao estudo de problemas relacionados à dinâmica de rotores; bem como, na operacionalização do sistema de aquisição de dados, condicionadores de sinais, excitadores, dispositivos de fixação de sensores e de dispositivos de segurança.

5.2 - Bancada de Dinâmica de Rotação

A bancada desenvolvida é composta de duas partes distintas: um rotor flexível, formado por elementos de eixo, disco e mancal, e uma estrutura fixa composta de duas chapas de aço espaçadas por colunas. Esta bancada apresenta um conjunto de características altamente desejáveis, tais como:

- Alta sensibilidade aos fenômenos inerentes à dinâmica de rotação devido à grande esbeltez apresentada pela árvore.
- Simplicidade de operação, devido à facilidade de fixação de instrumentos sensores e excitadores.
- Possibilidade de variação dos valores de rigidez e amortecimento dos mancais, de forma independente, em duas direções principais, mediante uma simples substituição de lâminas de diferentes espessuras e troca do óleo contido nos amortecedores.
- Facilidade de variação do posicionamento dos elementos de inércia ao longo do rotor, uma vez que estes são montados através de bucha autocentrante.
- Dispositivos limitadores de deslocamentos que propicia a segurança do operador bem como a integridade física do rotor e dos sensores.
- A rigidez da estrutura é bastante elevada quando comparada à rigidez do rotor.

O rotor é acionado por um motor de corrente alternada, trifásico, de 2 cv, modelo WEG 80/785. O motor é alimentado por um dispositivo variador de frequência (Micro Master Siemens 6SE3121 - 0cc40, 2200 W, 3 fases, 230 Volts), programável, que permite o ajuste da rotação desde zero a 120 Hz com uma resolução de 0.1 Hz.

A figura (5.1) mostra as características do rotor flexível. O eixo é construído em aço ABNT 1045, 17 mm de diâmetro e 1280 mm de comprimento, dos quais 1000 mm entre mancais,

retificado, polido e sobre o qual encontra-se montados 3 discos de 20 mm de espessura e de diâmetros 120 mm, 150 mm e 200 mm. Os discos podem ser posicionados livremente ao longo do eixo, sendo fixado por pinças metálicas que caracterizam engastes.

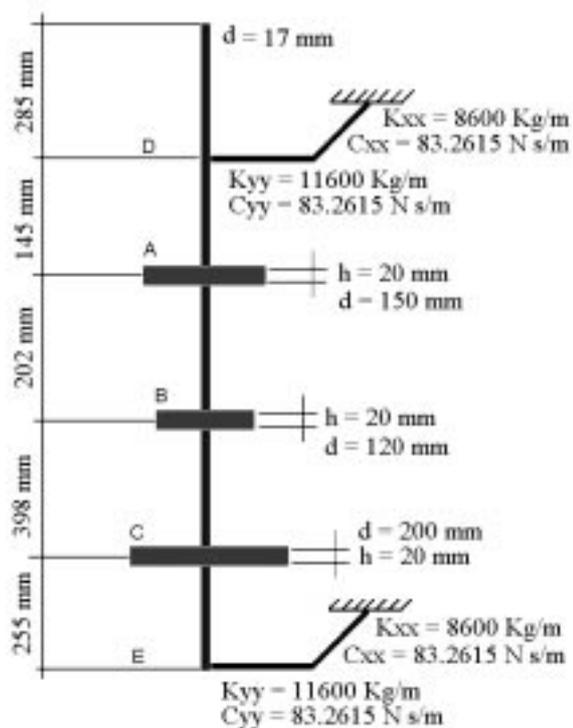


Figura (5.1) - Rotor: elemento móvel da bancada.

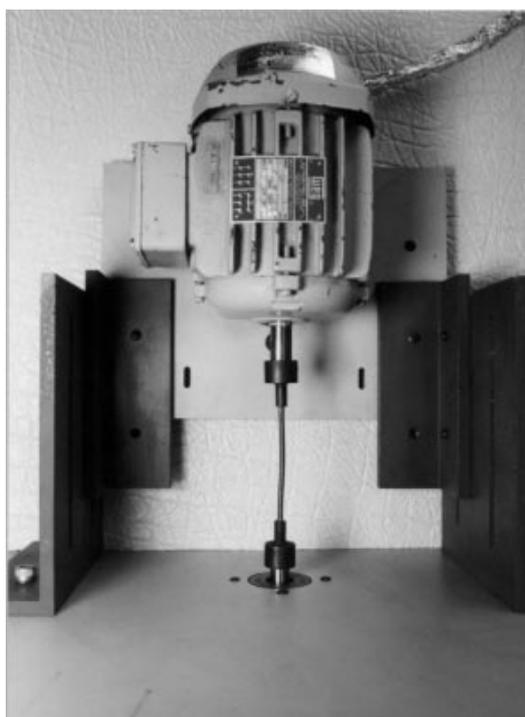


Figura (5.2) - Sistema de acionamento

A ligação entre o motor e o rotor, como mostrado na figura (5.2), é feita utilizando cabo de torção flexível, o que garante a transmissão apenas de torque entre o motor e a árvore do rotor. Outras características importantes desse acoplamento são a não introdução de rigidez adicional ao mancal e devido ao seu baixo peso e pequeno diâmetro, o efeito dinâmico introduzido na extremidade da árvore pode ser desprezado.

A suspensão do rotor é composto de lâminas, amortecedores e peças de ligação, como mostra a figura (5.3). A peça M1 (5.859 Kg) contém no seu interior, um rolamento autocompensador para a fixação da árvore e se move nas direções X e Y. A peça M2 (4.123 kg) move-se apenas na direção X. No programa de simulação do rotor, devido à característica do elemento utilizado para modelar a árvore, a direção X coincide com o centro da árvore, logo não segue o sistema de referência mostrado na figura (5.3).

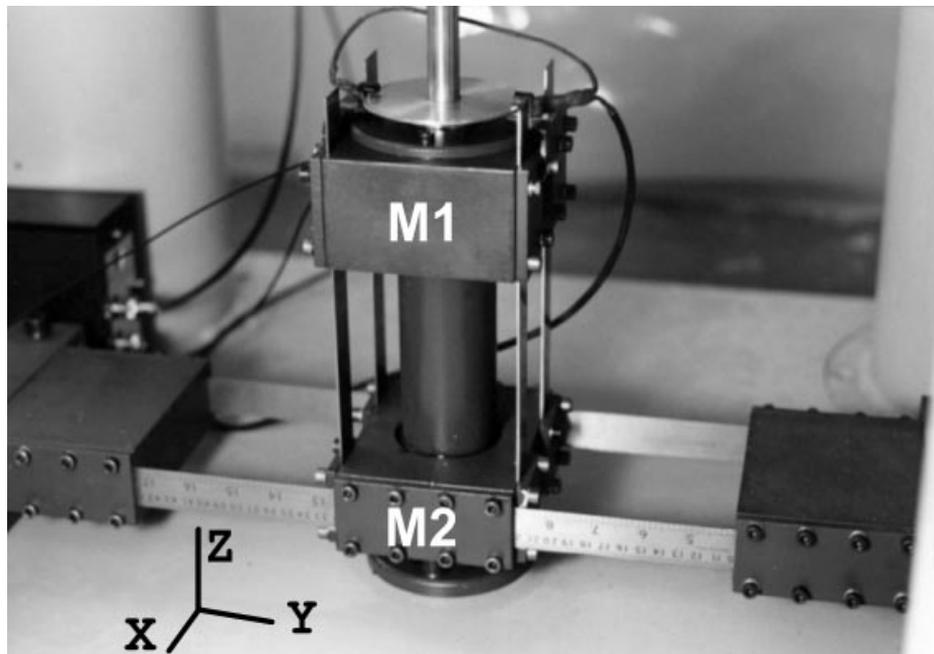


Figura (5.3) - Suspensão superior do rotor.

No mancal inferior foi instalado um encoder ótico com resolução de 3 graus. Existe uma posição de referência que permite disparar a aquisição dos sinais de vibração, sincronizada com uma posição angular do rotor.

A disposição dos amortecedores bem como o encoder de marcação da posição angular do rotor, pode ser visto na figura (5.5).

5.2.1 - Modelo computacional

A partir da concepção inicial do projeto da bancada, foi gerado um modelo computacional via elementos finitos (ANSYS) para cada uma das partes mencionadas. As quais foram analisadas separadamente.

É estrategicamente interessante que a estrutura que suporta o rotor tenha valores de frequências naturais de vibração tão afastadas quanto possível das frequências naturais do rotor, dentro da banda em que se espera operar do rotor. Esta condição propicia que os sinais medidos no rotor, não estejam substancialmente contaminados por vibrações oriundas da estrutura que o suporta.

A simulação numérica é uma ferramenta que auxilia o projetista na definição e alteração de características de projeto tais como espessura de chapas, forma do perfil e dimensão das colunas, materiais, etc., de modo a obter um comportamento desejado do equipamento a ser construído.

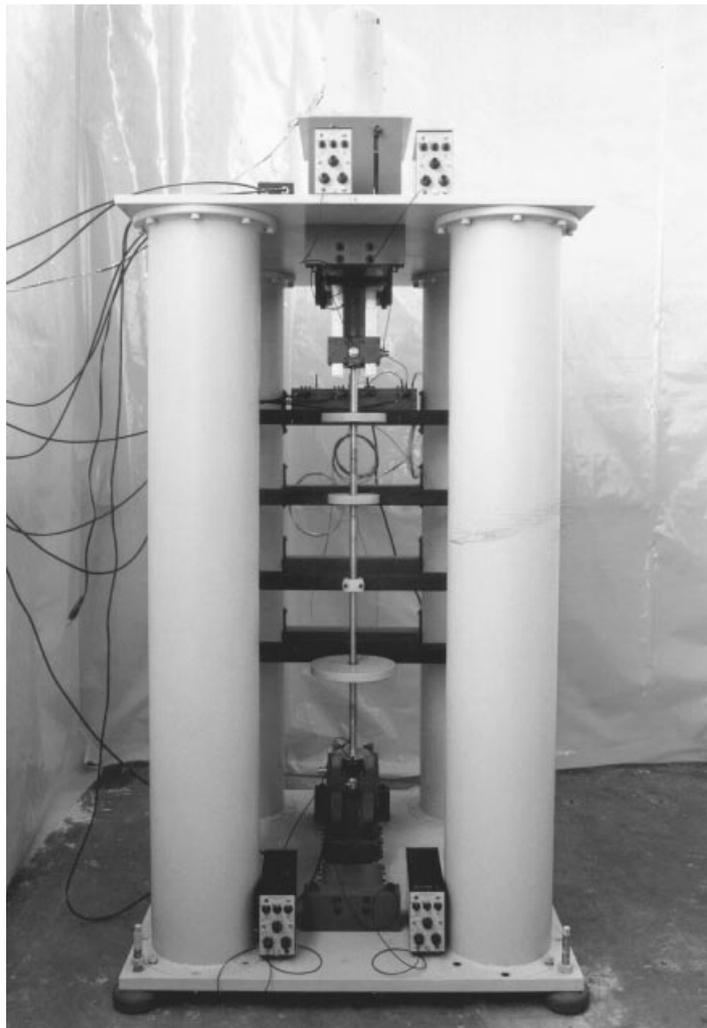


Figura (5.4) - Vista geral da bancada de teste de rotores

Após processo iterativo de alterações do modelo computacional, à luz dos resultados obtidos por simulação numérica, chegou-se a um projeto aceitável e realizável com base nas dimensões comerciais dos elementos envolvidos e processos de fabricação adequado às máquinas operatrizes disponíveis. No anexo I encontram-se os desenhos de todas as peças que compõem a bancada. Nos anexos II e III encontram-se as listagens dos modelos de elementos finitos utilizados para a simulação computacional da estrutura da bancada e do rotor.

A bancada foi montada sob sapatas flexíveis, como mostra figura (5.4). Tais sapatas estão disponíveis no mercado e sua aplicação principal é a sustentação de máquinas ferramentas. Tais elementos permitem o nivelamento da máquina ferramenta e atuam como suspensão que filtram as vibrações mecânicas de frequências altas, propagadas através do solo.

5.2.2 - Caracterização experimental do rotor

A caracterização da rigidez dos mancais foi feita através dos métodos analítico e experimental. Como as molas são lâminas, a formulação analítica é baseada na teoria de flexão de viga biengastada.

Pelo método estático, a rigidez é obtida utilizando-se de diferentes valores de forças para provocar deflexões nas molas. Estas deflexões são então monitoradas por um relógio comparador. A rigidez da mola é dado pelo coeficiente angular da reta que interpola os pontos de força versus deslocamento.

Foi realizado também a caracterização da rigidez através do método dinâmico, que consiste em analisar o conteúdo em frequência da resposta do mancal ω_n quando excitado por uma força impulsiva, antes e depois da adição de uma massa de tara m_{tara} conhecida. Através da variação da frequência é possível determinar a rigidez do mancal K_{mancal} , uma vez que:

$$\omega_n = \sqrt{\frac{K_{mancal}}{M_{mancal} + m_{tara}}} \quad (5.1)$$

Assis (Assis, 1999a) obteve numericamente os valores de rigidez, utilizando a técnica dos algoritmos genéticos. Os valores encontrados estão sumarizados na tabela (5.1). Os mancais inferior e superior são idênticos.

Tabela (5.1) - Valores de rigidezes dos mancais.

	Analítico	Estático	Dinâmico	Algoritmo Genético
Kxx	8677.50 N/m	8601.00 N/m	8593.00 N/m	9806.00 N/m
Kyy	11666.00 N/m	11610.00 N/m	11598.00 N/m	11521.00 N/m

Para a determinação dos coeficientes de amortecimento foi feita uma montagem experimental com uma mesa de um grau de liberdade. Esta mesa foi excitada através de condição inicial de deslocamento. A resposta da aceleração do sistema amortecido foi então digitalizada (ver anexo IV). Conhecida a massa da mesa (23.376 Kg), o período de oscilação (0.247 segundos) e decremento logaritimico da curva de resposta, o coeficiente de amortecimento foi determinado.



Figura (5.5) - Vista da configuração da suspensão inferior do rotor.

Todos os amortecedores da bancada contém a mesma geometria e o mesmo tipo de óleo (Fluído Hidráulico TM 30).

Assis (Assis, 1999b) calculou numericamente o coeficiente de amortecimento utilizando algoritmos genéticos.

Tabela (5.2) - Valores do coeficiente de amortecimento dos mancais.

	Decremento Logarítimo	Algoritmo Genético
Cxx	83.1037 N.s/m	81.8498 N.s/m
Cyy		88.4477 N.s/m

O diagrama de Campbell característico do rotor, mostrado na figura (5.6), foi mapeado excitando alternadamente os discos A e B, como mostrado pela figura (5.1), através de uma força impulsiva aplicada pela ação de um pêndulo. Na posição dos discos e mancais foram posicionados sensores de proximidade. O sinal de resposta em frequência desses sensores foram calculados para um número de médias não inferior a 30. Este procedimento foi executado na direção X e Y, para o rotor girando desde zero a 63 Hz para os seguintes valores de frequências: 0, 3.41, 10.52, 15.47, 24.81, 30.43, 38.90, 43.68, 52.35, 57.41 e 63.00 Hz.

No anexo V pode ser visto o gráfico da função resposta em frequência do rotor para a rotação nula.

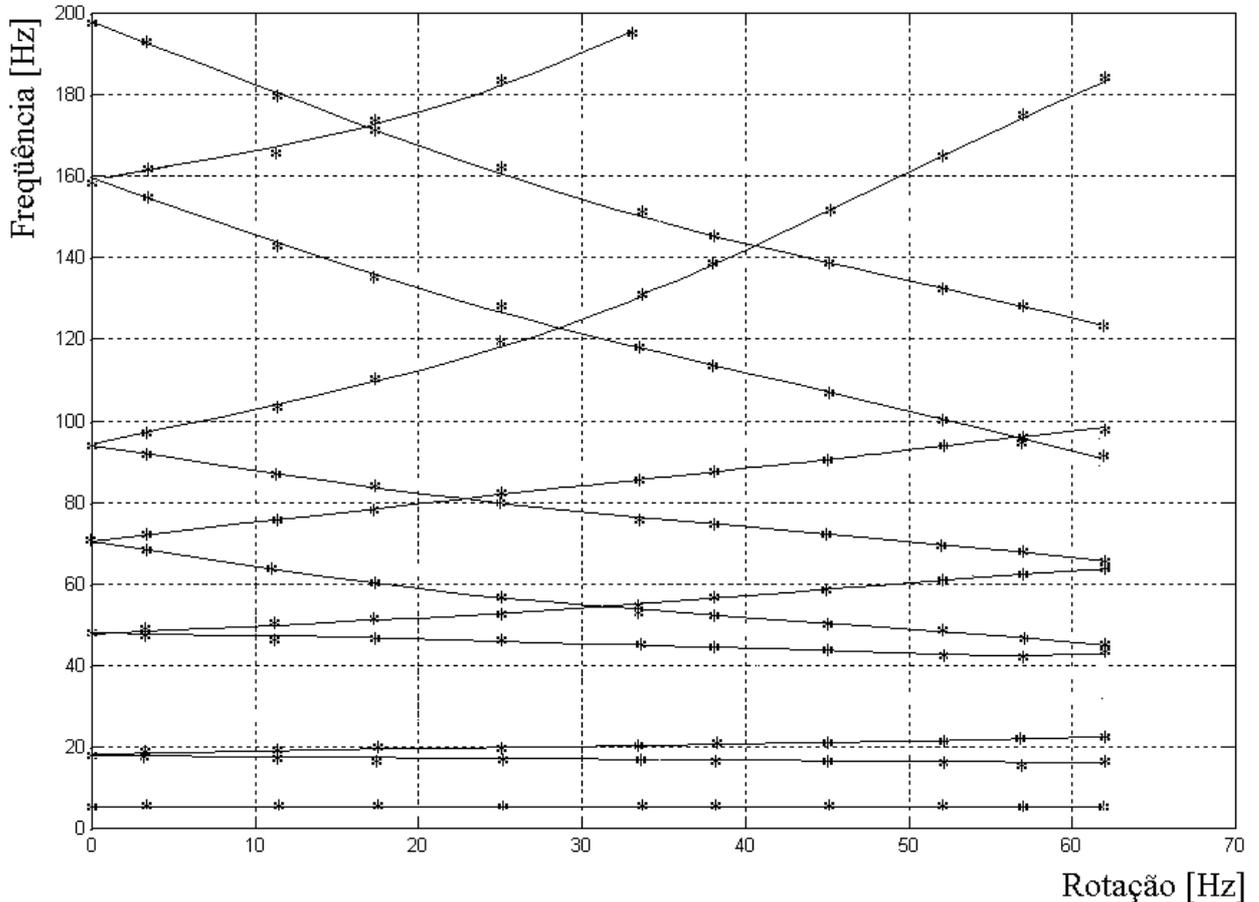


Figura (5.6) - Diagrama de Campbell do rotor, obtido experimentalmente.

5.2.3 - Caracterização experimental da estrutura

Após a construção e montagem da estrutura, com o auxílio de um analisador espectral, de um martelo equipado com célula de carga e um acelerômetro, foram obtidos os gráficos das funções de transferência e coerência, como mostrado no anexo VI, entre os pontos A_1A_2 , B_1B_2 e C_1C_2 , como mostra a figura (5.7). O índice 1 refere-se ao ponto da aplicação da força impulsiva e o índice 2 é o ponto de medição do acelerômetro. Esses gráficos permitem a identificação das frequências de corpo rígido, de torção e de flexão nas direções Y e X. A frequência de corpo rígido é da ordem de 11.88 Hz.

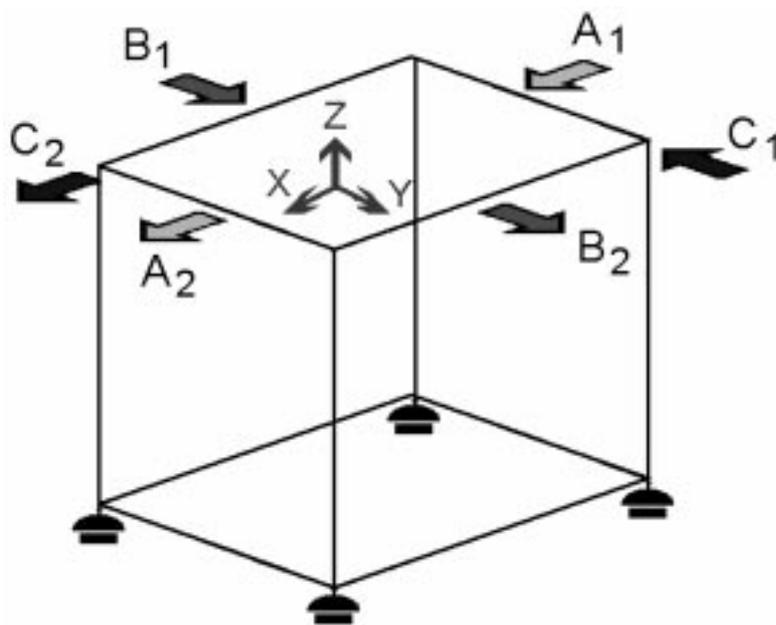


Figura (5.7) - Ensaio das freqüências naturais da estrutura

5.2.4 - Comparações de resultados

Para o rotor, a tabela (5.3) abaixo mostra a comparação entre as freqüências naturais calculadas numericamente e as obtidas de forma experimental para rotação nula.

Tabela (5.3) - Freqüências naturais do rotor para $\Omega = 0$.

Freqüências do rotor [Hz]	
Experimental	ANSYS
5.25	5.21
19.37	19.06
46.37	48.11
73.25	64.88
158.50	171.76
197.50	217.45

A figura (5.8) mostra alguns dos primeiros modos de vibrar do rotor no plano XZ obtidos através do ANSYS. Os resultados no plano YZ são semelhantes.

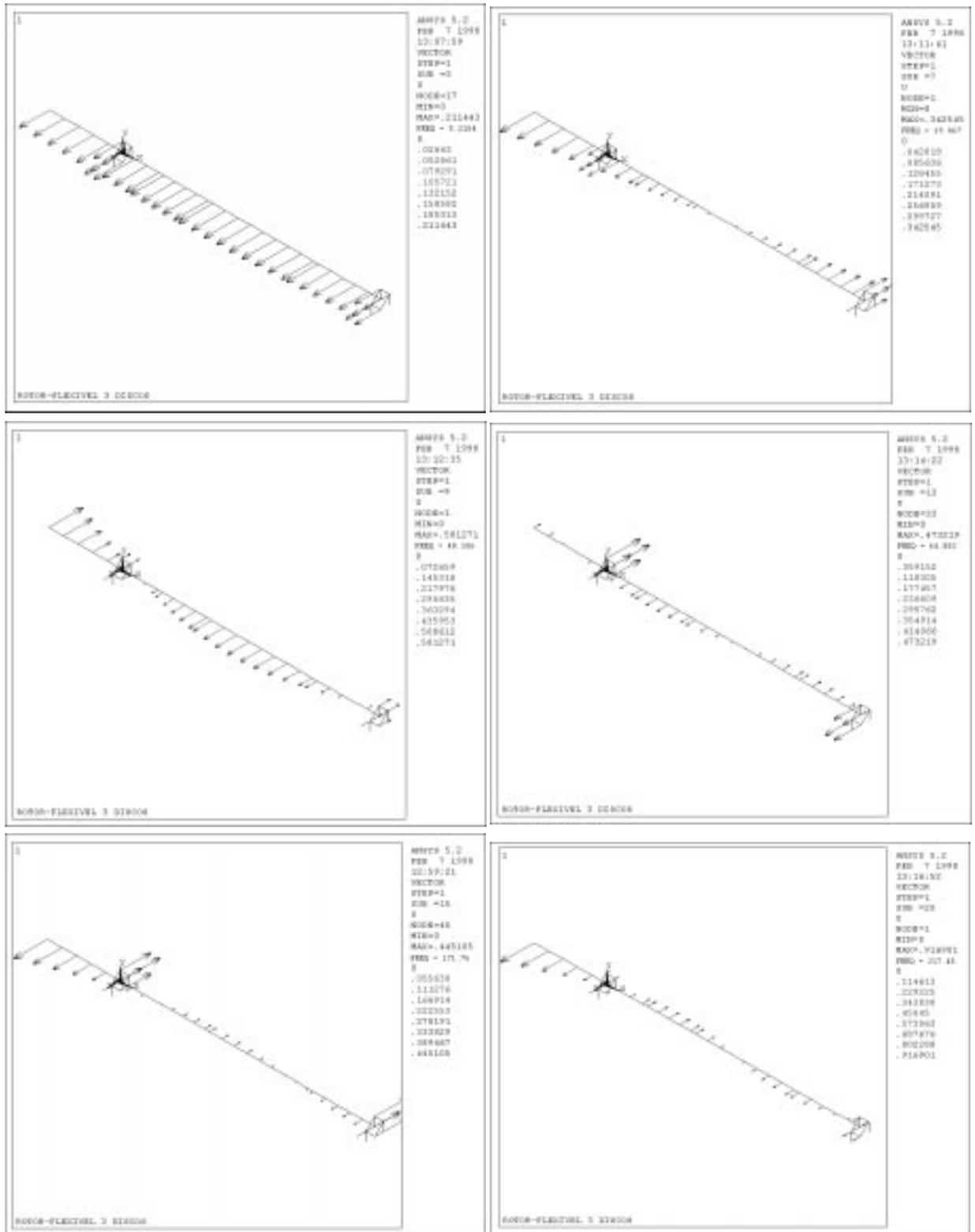


Figura (5.8) - Modos de vibrar do rotor.

De cima para baixo, da esquerda para direita tem-se: 1º modo de corpo rígido, 2º modo de corpo rígido, 1º modo de flexão, 2º modo de flexão, 3º modo de flexão, 4º modo de flexão.

Na figura (5.9) são mostrados as formas dos seis primeiros modos de flexão da estrutura também obtidos numericamente.

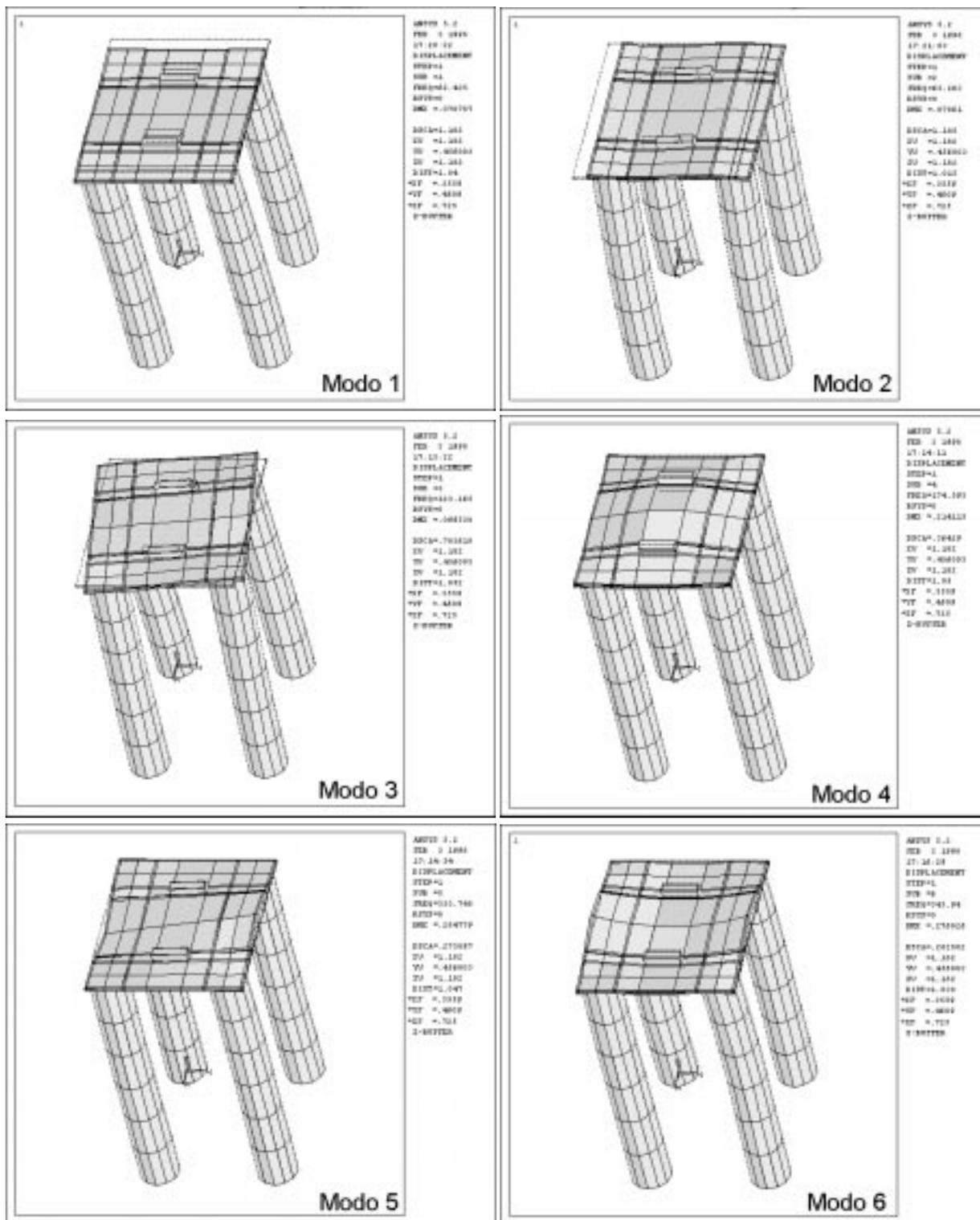


Figura (5.9) - Primeiros modos de flexão da estrutura.

Para a estrutura fixa a tabela (5.4) mostra uma comparação entre as frequências naturais

de vibração calculadas pelo ANSYS e experimentalmente. Pode-se observar satisfatória concordância entre os resultados. A formulação do modelo computacional não prevê a obtenção das freqüência de corpo rígido da estrutura, uma vez que as colunas foram consideradas engastadas na placa inferior e esta solidária ao solo.

Tabela (5.4) - Freqüências naturais da estrutura.

Freqüências da Estrutura [Hz]	
Computacional [ANSYS]	Experimental
----	11.50
52.44	55.00
62.25	69.50
123.16	104.75
174.59	186.25
333.75	311.25
343.94	452.50

O modelo computacional não prevê o acréscimo de rigidez provocado pela fixação do motor, por esse motivo, nas altas freqüências, os resultados do modelo experimental se distanciam mais dos resultados do modelo computacional.

As travessas que suportam os amortecedores, assinalada com a letra 'A' na figura (5.5), possuem freqüência de vibração lateral de 244.0 Hz. As travessas que suportam os sensores de proximidade, assinaladas com as letras 'B' e 'C', na figura (5.10), possuem freqüências naturais de vibração lateral de 159.0 Hz e 364.0 Hz respectivamente, para os primeiros modos de vibração

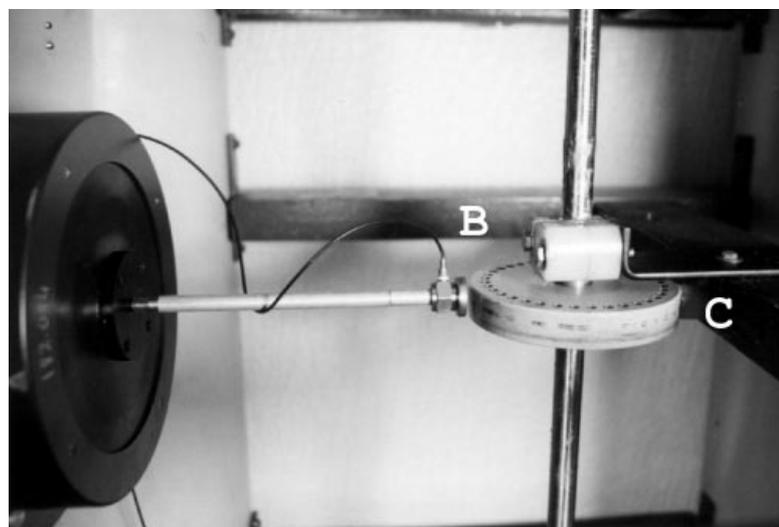


Figura (5.10) - Travessas de suporte dos sensores de proximidade.

A rotação de operação do rotor foi selecionada observando uma faixa de freqüência livre tanto no diagrama de Campbell, figura (5.6), como nos gráficos das funções de transferência da estrutura, como por exemplo a figura (5.11), que mostra a função de transferência da estrutura, excitada no ponto A1 e medido no ponto A2, como indicado pela figura (5.7).

Da figura (5.11) observa-se a existência de uma faixa de freqüência disponível, que vai desde a freqüência de corpo rígido, 11.500 Hz até a primeira freqüência de flexão 50.00 Hz. Do diagrama de Campbell, tem-se uma faixa freqüência livre entre 20.00 Hz e 40.00 Hz, para a rotação desde zero a 63.00 Hz. A rotação selecionada para operação do rotor foi de 35.00 Hz.

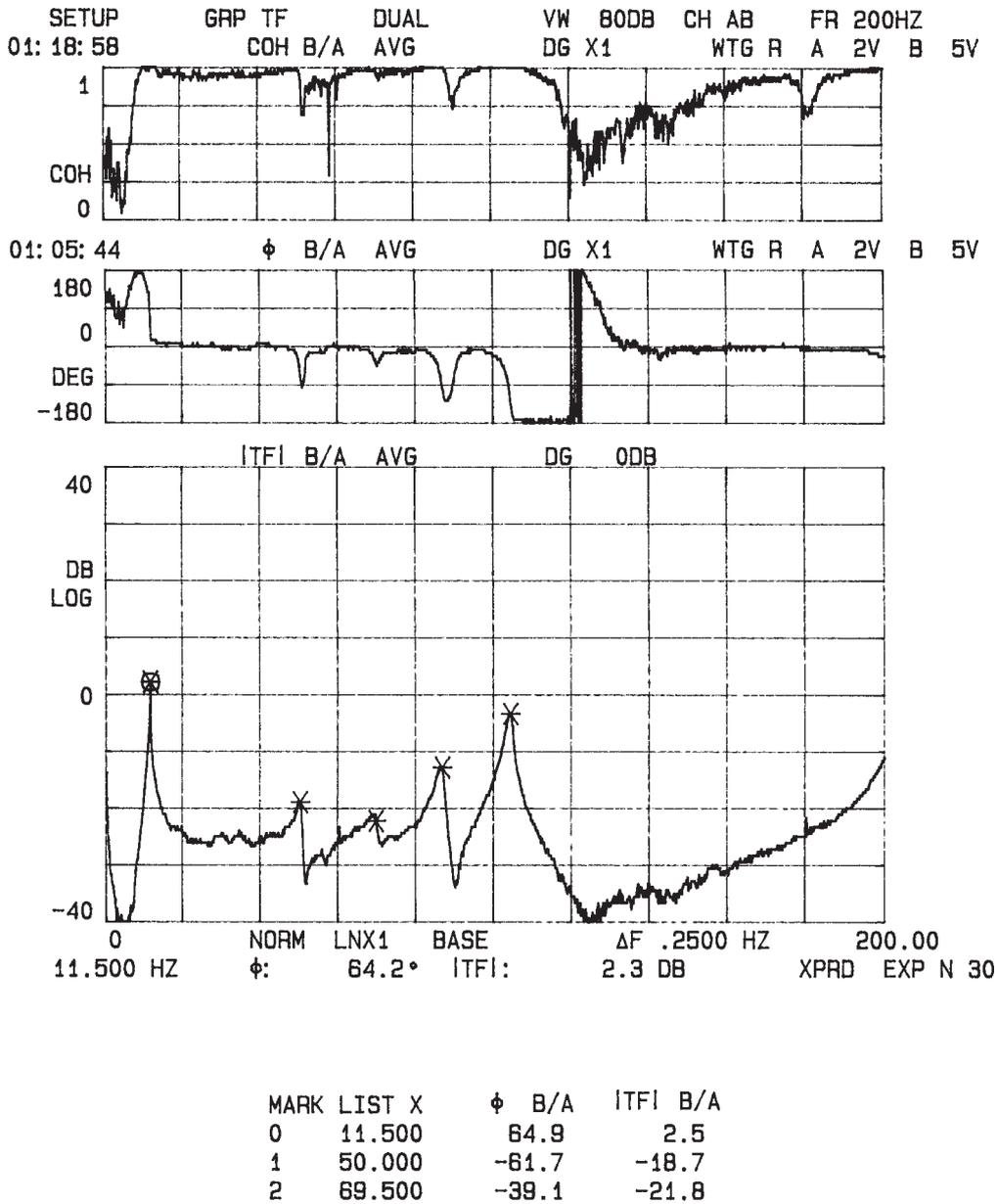


Figura (5.11) - Função de transferência da estrutura, excitação: A1 e medição: A2.

5.3- Sistema de Medição

O sistema de digitalização e aquisição de dados, mostrado pela figura (5.13), é composto de uma estação Risc modelo HP Série 700i com impressora HP LaserJet 5p; fonte DC programável modelo HP 35653A; sistema modular de aquisição e processamento de sinais HP 35650, com os seguintes módulos: gerador de sinais HP 35653C (102.4 KHz), tacômetro/trigger HP 35658A, modulo HP 35655A Input/Output (8 canais simultâneos, freqüência máxima de análise de 12.8 KHz, cada canal equipado com filtro passa baixo com freqüência de corte ajustado para freqüência de análise), processador de sinais HP 35654B. A figura (5.12) mostra o diagrama de blocos do sistema de medição.

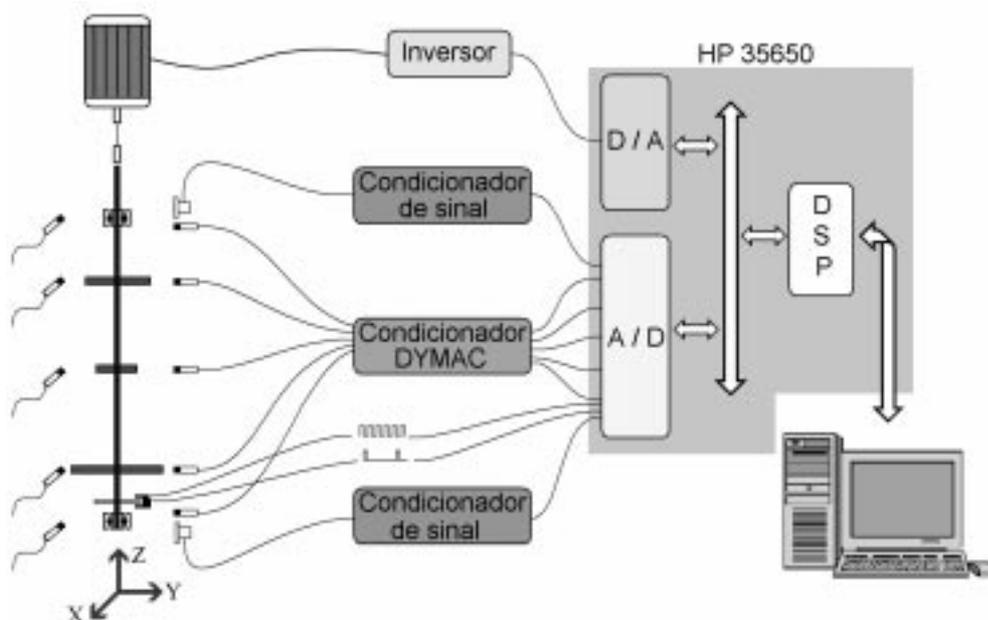


Figura (5.12) - Diagrama de blocos do sistema de medição.

Na figura (5.12), a rotação do motor é ajustada pelo inversor. O inversor é comandado remotamente por um nível de tensão DC disponibilizado pelo conversor digital/analógico do HP35653A.

A estrutura da bancada desenvolvida permite a fixação de sensores de proximidade ao longo do rotor nas direções X e Y, bem como a fixação de acelerômetros nos mancais. No mancal inferior do rotor encontra-se montado o encoder ótico que produz dois sinais, um sinal de sincronismo caracterizado por um único pulso a cada volta do rotor e um sinal de monitoramento da posição angular com 120 pulsos por volta do rotor.

Os sinais dos sensores são digitalizados pelo conversor analógico/digital, filtrados por um filtro passa baixo (freqüência de corte = freqüência de análise) localizado na entrada de cada um dos canais, e posteriormente enviados a central de processamento (DSP) onde é calculada

a transformada de Fourier (FFT), funções de transferência, etc.. Depois que o DSP realizou os cálculos estipulados pelo usuário, os dados processados são enviados à estação HP 700i, onde são plotados os gráficos e gerados os arquivos em disco rígido.



Figura (5.13) - Sistema de aquisição de dados

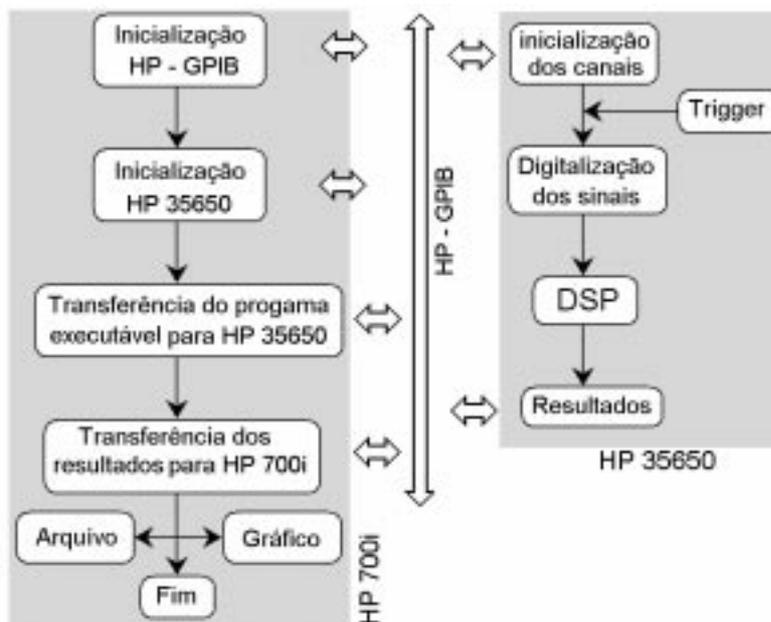


Figura (5.14) - Diagrama de blocos do programa de aquisição de dados.

A figura (5.14) mostra o diagrama de blocos típico do programa de aquisição de dados do HP 35650. A estação HP 700i está ligada ao módulo de processamento e aquisição de dados

HP 35650 através de uma interface HP-GPIB.

São dois os programas executáveis necessários para a operação do sistema de aquisição e processamento de dados HP. O programa principal é executado pela estação de trabalho HP 700i e o segundo programa é executado pelo módulo HP 35650.

O programa principal é responsável pelas seguintes operações: inicialização da interface HP-GPIB, inicialização do módulo HP 35650, indicação dos dados solicitados pelo usuário, transferência do segundo programa no módulo HP 35650, enviar um sinal de início para o programa secundário, transferir os dados processados pelo DSP (CPU dedicada ao processamento de sinais) para a estação HP 700i, desenho dos gráficos dos dados solicitados na tela e armazenamento dos dados em arquivo.

O programa executado pelo módulo HP 35650, uma vez colocado em funcionamento, ajusta a frequência de corte dos filtros passa baixo (um para cada canal) e espera por um sinal de disparo para proceder a leitura dos dados armazenados na memória de cada canal. Os dados lidos são transferidos para o DSP. Os sinais processados pelo DSP são movidos para uma área de memória local que é acessada pela estação HP 700i.

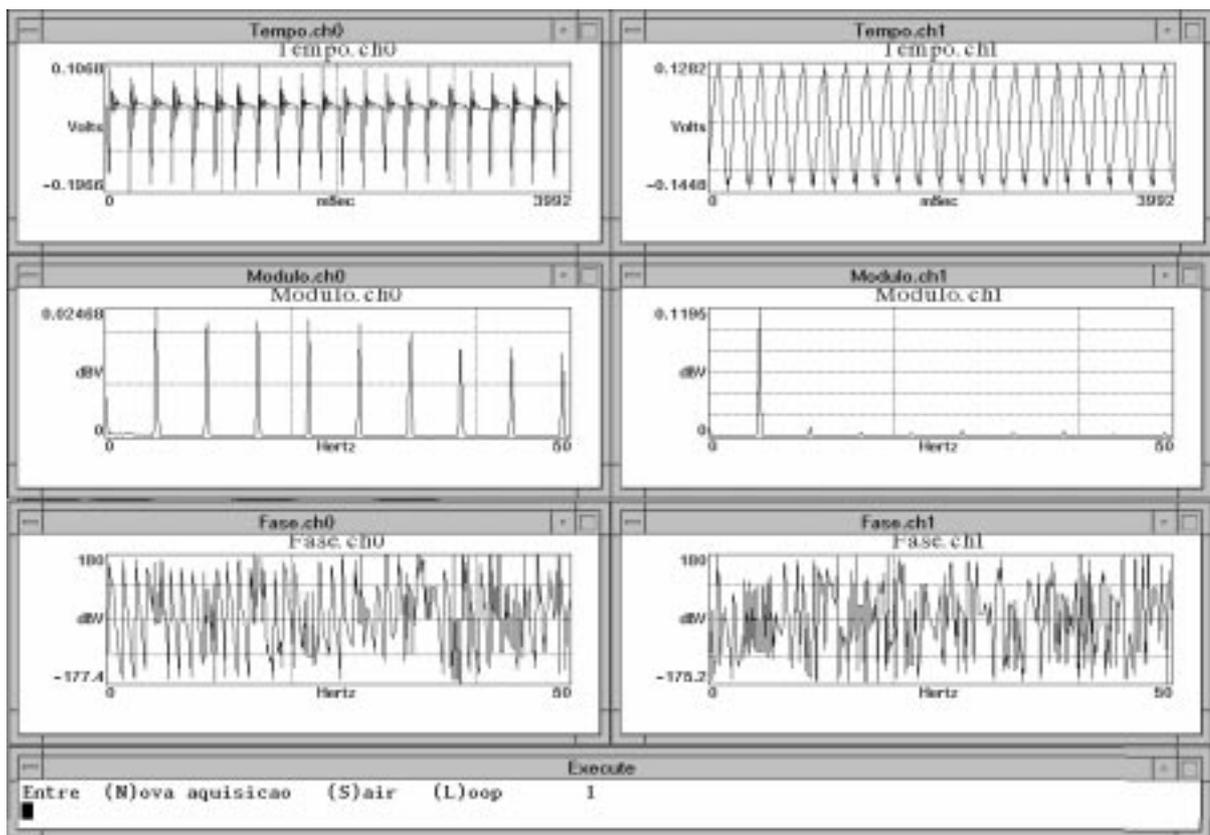


Figura (5.15) - Interface gráfica do programa de aquisição.

Foram desenvolvidos dois programas que operam o sistema de aquisição de dados HP, utilizando as funções básicas do hardware fornecidas pelo fabricante (MIL - Measurement Interface Library). O primeiro programa adquire, gera arquivo dos dados e plota os sinais no domínio do

tempo e da frequência, o sinal de todos os canais. A figura (5.15) mostra a interface desse programa. Um segundo programa, adquire e obtém a função de transferência de cada canal com relação ao canal zero.

A instrumentação utilizada no funcionamento da bancada são: acelerômetros piezoelétricos (Bruel & Jjær, tipo 4367), célula da carga piezoelétrica (Bruel & Jjær, tipo 8200), condicionadores de sinais (Bruel & Jjær, tipo 2635), sensores de deslocamento eletromagnéticos (DYMAC modelo M60) e um encoder ótico (construído) de monitoração da posição angular do rotor e disparo do trigger do sistema de aquisição de dados.

Foi desenvolvido um programa que, a partir da base de dados do ANSYS, extrai as matrizes elementares e monta as matrizes globais de massa, rigidez, amortecimento (giroscópica) e o vetor de força generalizado, em seguida, elimina os graus de liberdade não utilizados. De posse das matrizes. Foi também codificado um programa que permite a simulação numérica da resposta do rotor. A listagem desses programas encontram-se no anexo VIII.

Foi necessário a construção de um circuito composto de um conjunto de amplificadores inversores, com alta impedância na sua entrada. Este circuito permite ajustar o nível de tensão (ganho) e faixa de variação (offset) dos sinais advindos dos sensores de proximidade de modo a torna-los compatíveis com às características do modulo HP 35655A. A figura (5.16) mostra o esquema elétrico de um dos amplificadores. No anexo VII encontram-se as curvas de calibração dos sensores de proximidade.

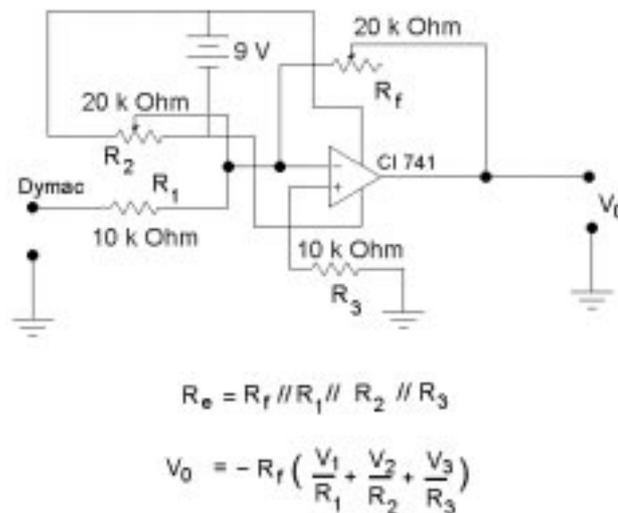


Figura (5.16) - Circuito amplificador inversor.

Em todas as medições foram usados os seguintes ajustes: frequência de análise: 50 Hz, frequência de amostragem: 128 Hz (2.56 vezes 50 Hz, para proteção anti-aliasing). Filtros dos amplificadores de carga ajustados para: frequência de corte inferior de 2 Hz e frequência de corte superior de 100 Hz.

Capítulo 6

Fontes de Excitação

O objetivo aqui é desenvolver uma metodologia baseada em redes neurais, capaz de identificar algumas das principais fontes de excitação que afetam as máquinas rotativas.

Três tipos de excitações foram estudados: excitação por desbalanceamento, excitação por força assíncrona de natureza magnética aplicada no disco e excitação assíncrona aplicada nos mancais. Tais excitações caracterizam excitações provenientes de máquinas elétricas, mancais magnéticos, defeitos em mancais e excitações externas.

As respostas do sistema às excitações foram adquiridas através de sensores de proximidade magnéticos (DYMAC), ajustados para medir a variação do deslocamento. Como mostra a figura (6.1), os sensores foram numerados e posicionados num mesmo plano, ao longo do rotor, em cinco posições: nos três discos e no mancal superior e inferior.

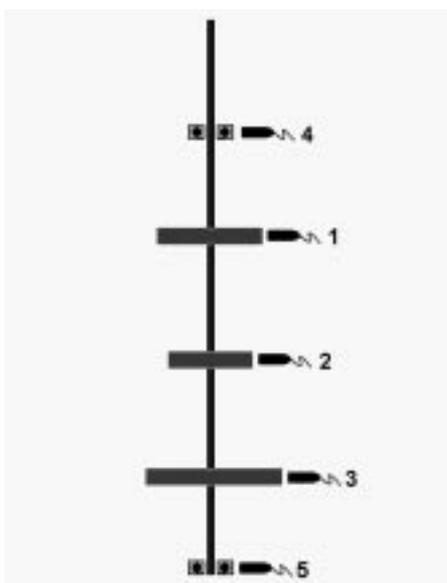


Figura (6.1) - Posicionamento dos sensores.

Um sinal advindo do rotor é submetido à uma rede neural classificadora, que determina se o sinal é característico do sistema não perturbado, caso contrário, a natureza da fonte excitadora será indicada pelo neurônio da camada de saída, com o maior nível de ativação.

Para cada um dos tipos de excitação foi treinada um rede neural específica, que determina a sua localização e severidade.

Determinada a natureza da excitação, o sinal é condicionado para ser submetido à rede neural treinada para determinar a localização e a quantificação da excitação.

6.1 - Excitação por desbalanceamento.

Como mostrado pela figura (6.1), a bancada possui três discos. Cada um dos discos é dotado de 36 furos circularmente espaçados, tal como mostra a figura (6.2).

A rotação de regime, 35 Hz, foi arbitrariamente escolhida por estar situada numa região segura de operação, ou seja, não muito próxima de ressonâncias, como mostrou o diagrama de Campbell, figura (6.9). É claro que, numa aplicação real, a rotação de regime indicada pelo fabricante, para uma máquina rotativa, sempre é uma rotação segura de funcionamento, uma vez que esta é uma das principais preocupações na fase de projeto.

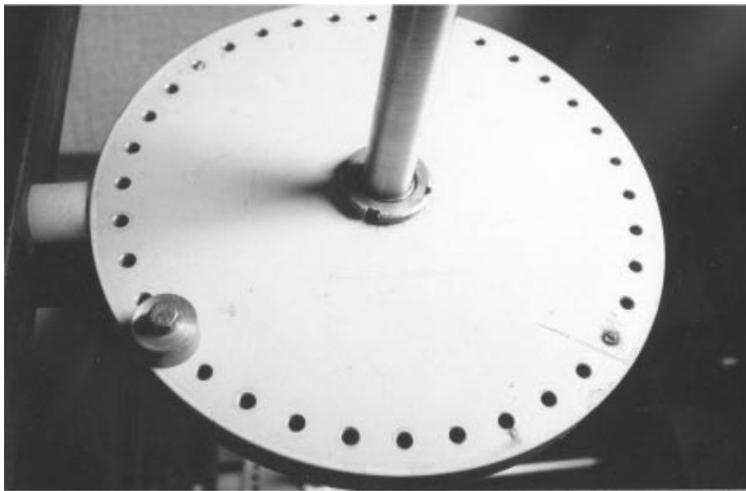


Figura (6.2) - Posicionamento das massas de desbalanceamento

Com o auxílio de diferentes massas de desbalanceamento, foram adquiridos os sinais da resposta do sistema, simultaneamente pelos cinco sensores, sempre para a mesma rotação de funcionamento.

O conjunto das massas de desbalanceamentos continham valores tais que as respostas do sistema apresentavam desde pequenos até grandes deslocamentos; sem contudo atingir os batentes de segurança.

Os dados adquiridos foram obtidos variando o valor do desbalanceamento, o disco e o furo (ângulo com relação à marca de disparo do sistema de aquisição de dados). Foram realizados também testes com desbalanceamentos simultâneos em mais de um disco.

A frequência de análise foi de 50 Hz (2.56 vezes 50 Hz, para proteção anti-aliasing). Frequência de corte inferior de 2 Hz e superior de 100 Hz, ajustados nos amplificadores de cargas.

Foram realizados 60 experimentos. 30 médias foram computadas para cada uma das medidas adquiridas.

6.2 - Excitação por força de natureza magnética.

As excitações de natureza magnética foram aplicadas ao disco intermediário conforme a figura (6.3). Com o auxílio de um gerador de sinais, um sinal senoidal foi amplificado e fornecido a um excitador (shake) acoplado a uma base magnética. Esta, por sua vez exercia força de atração sobre disco metálico.

O sinal da força aplicada ao rotor foi monitorado por uma célula de carga situada entre a base magnética e o excitador, conforme ilustra a figura (6.3).

Foram sintetizados sinais de diversos valores com frequência menores que 50 Hz. Para cada frequência a amplitude foi ajustada de tal sorte causar a maior perturbação possível no rotor, sem contudo permitir a colisão entre o disco e a base magnética.

Durante todos os ensaios o rotor foi mantido em rotação de regime, 35 Hz. A frequência de análise foi de 50 Hz (2.56 vezes 50 Hz, para proteção anti-aliasing). Frequência de corte inferior de 2 Hz e superior de 100 Hz, ajustados nos amplificadores de cargas. Número de médias para cada medida adquirida foi ajustado para 100.

O plano de excitação é o mesmo plano dos sensores, porém em direção oposta a estes.

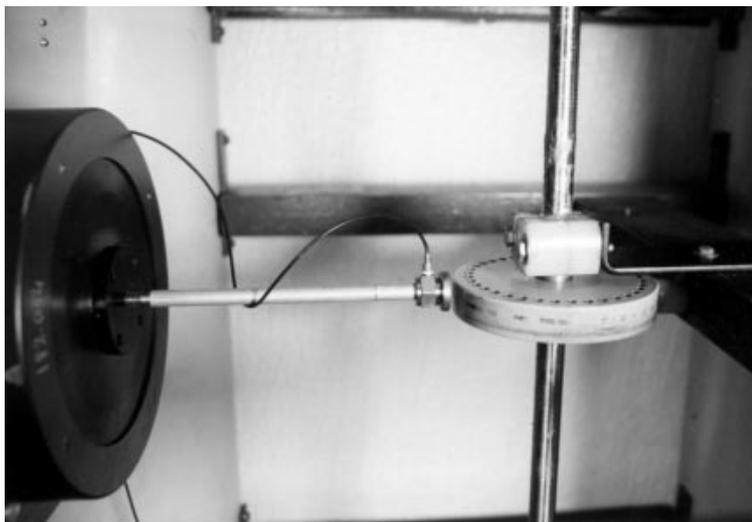


Figura (6.3) - Detalhe da excitação de natureza magnética.

Excitação por forças com características magnéticas é uma excitação periódica com substancial quantidade de energia em torno da frequência principal de excitação. A figura (6.4) mostra um sinal característico da força desse tipo de excitação. Foram realizados 60 experimentos dentre os quais 20 foram reservados para a qualificação da rede neural após o treinamento.

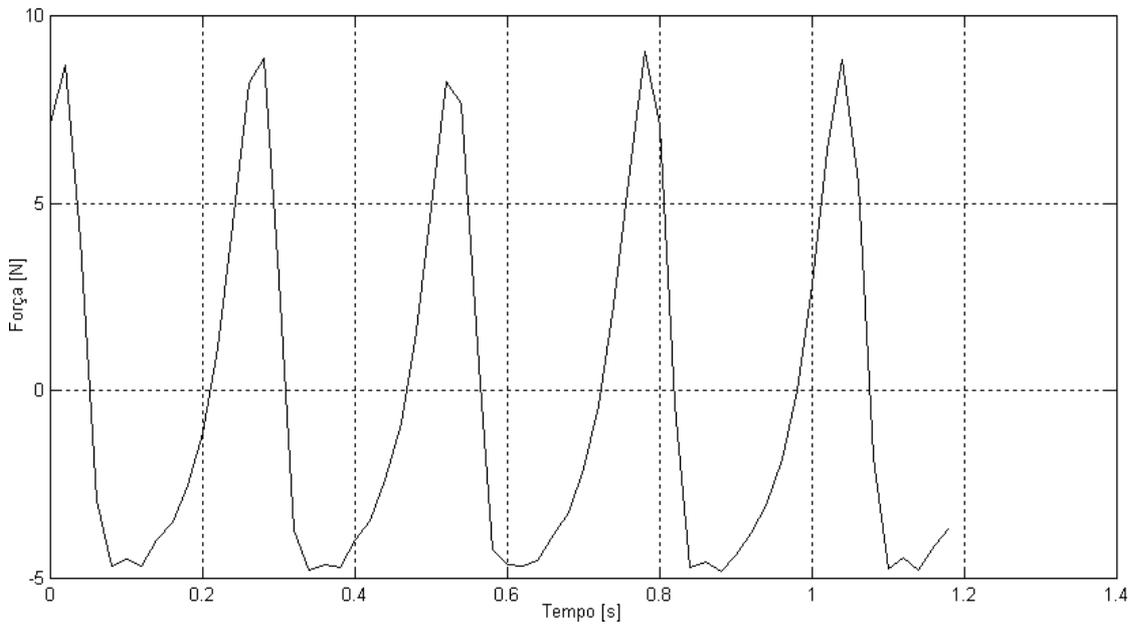


Figura (6.4) - Força magnética aplicada ao disco 2, amplitude 9 N, frequência 10 Hz.

Em cada experimento foram digitalizados 512 pontos no domínio do tempo e 401 pontos no domínio da frequência, devido à proteção anti-alias (Hewlett-Packard, 1993). Os 111 pontos restantes são os pontos pertencentes à chamada banda de guarda.

A biblioteca MIL (Measurement Interface Library), do sistema de aquisição de dados da HP, utiliza a seguinte expressão para o cálculo da largura de banda:

$$\text{Banda de guarda} = \left(\frac{7}{32} \times f_{max} \right) - 1 \tag{6.1}$$

onde f_{max} , múltiplo de 2^n , é a frequência máxima.

O uso de toda a banda amostrada, 401 pontos de magnitude e 401 pontos de fase, ou seja, 802 variáveis de entrada para a rede neural, inviabiliza o uso dessa técnica para esta aplicação. Como já fora discutido na seção 4.9, nesta aplicação é evidente a presença de variáveis redundantes, ou seja, linearmente dependentes (frequências naturais e seus harmônicos), bem como a presença de variáveis não correlacionadas (ruído). Outra consequência natural, advinda do número elevado de variáveis, é a grande demanda de memória e o alto custo computacional necessário ao treinamento da rede neural.

A metodologia descrita na seção 4.9 foi utilizada para promover a redução do número de variáveis de cada experimento. Diferentes níveis de redução foram implementados. A sua influência será posteriormente discutida.

6.3 - Excitação dos mancais por força assíncrona.

Excitações assíncronas foram aplicadas nos mancais, um de cada vez, no mesmo plano dos sensores porém na direção oposta a estes.

Sinais senoidais sintetizados, amplificados foram fornecidos a um excitador (shake) acoplado ao mancal por meio de uma haste flexível. Entre a haste e mancal foi fixado uma célula de carga piezoelétrica, através da qual o sinal da força aplicada ao rotor foi monitorado. A figura (6.5) mostra a montagem do excitador atuando no mancal superior da bancada.

Mantida a rotação de regime, 35 Hz, foram adquiridas as respostas para diversos valores de frequência (menores que 50 Hz) e amplitude.

A frequência de análise foi de 50 Hz (2.56 vezes 50 Hz, para proteção anti-aliasing) (Hewlett-Packard, 1993). Frequência de corte inferior de 2 Hz e superior de 100 Hz, ajustados nos amplificadores de cargas. Número de médias de aquisição ajustado para 30.

O sinal no domínio da frequência, proveniente deste tipo de excitação, é um sinal limpo, com energia concentrada em poucos picos. É fácil identificação visual das frequências de excitação, de rotação bem como os seus harmônicos. Foram realizados 60 experimentos dentre os quais 20 foram reservados para a qualificação da rede neural após o treinamento.

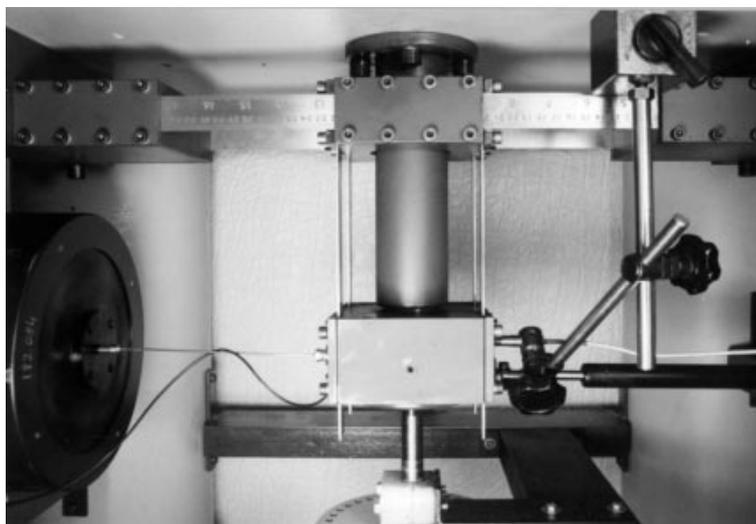


Figura (6.5) - Detalhe da excitação assíncrona nos mancais.

Capítulo 7

Teste da Metodologia Proposta

7.1 - Introdução.

No capítulo 1, seção 1.3, foi descrita, a metodologia com qual se pretende localizar e caracterizar fontes de excitação num sistema mecânico.

Foi implementado em linguagem C, um programa de rede neural de treinamento supervisionado baseada na teoria de back-propagation, como descrita no capítulo 4. O algoritmo de minimização dos erros foi implementado usando o método dos gradientes descendentes, com passo variável e memória do passo anterior (momentum). A listagem deste programa está no anexo IX.

Foi implementado em linguagem MATLAB, um programa de compactação de dados utilizando a técnica estatística de seleção das variâncias principais, tal como descrita no capítulo 4, seção 4.9. A listagem deste programa encontra-se no anexo X.

O objetivo principal desse trabalho é localizar e caracterizar três tipos diferentes de excitação em máquinas rotativas: excitação por desbalanceamento, excitação assíncrona localizada nos mancais e excitação assíncrona de característica não linear, aplicada num dado ponto do rotor. A primeira das redes neurais deverá distinguir as três classes de excitação bem como o estado não excitado do rotor. As demais redes neurais são redes neurais particularizadas, uma para cada tipo de excitação. Essas redes neurais deverão fornecer a localização e os parâmetros que caracterizam as excitações para as quais foram treinadas.

O diagrama de blocos abaixo mostra a seqüência de procedimentos a ser adotado.

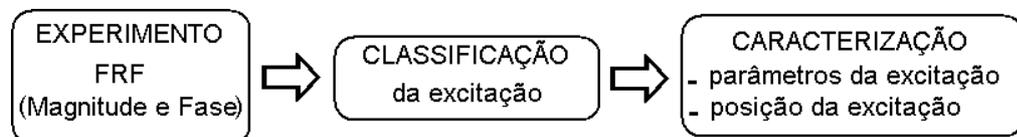


Figura (7.1) - Fluxograma simplificado do procedimento de caracterização das excitações

7.2 - Modelo de 3 Graus de Liberdade

Como objetivo de validar o programa de rede neural implementado, usou-se um código implementado em MATLAB, representativo de um sistema dinâmico de três graus de liberdade

como mostra a figura (7.2).

Foi gerado um banco de dados formado pela resposta em frequência do sistema (FRF), calculados variando-se a amplitude da força entre 1.0 e 10 N, bem como o seu ponto de ação (m_1 , m_2 e m_3). As FRFs foram obtidas para 100 pontos dentro da banda que contém as frequências naturais do sistema.

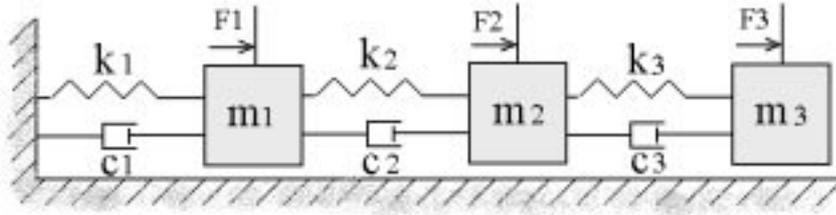


Figura (7.2) - Sistema mecânico de 3 graus de liberdade (gdl).

A figura (7.3) mostra as funções resposta em frequência do sistema 3 gdl, medido nas massas m_1 , m_2 e m_3 , para os seguintes valores: $m_1=1.0$ Kg, $m_2=2.0$ Kg, $m_3=3.0$ Kg, $k_1=10.0$ N/m, $k_2=60.0$ N/m, $k_3=70.0$ N/m, $k_4=10.0$ N/m, $c_1=c_2=c_3=c_4=0$, $f_1=f_3=0$ e $f_2=10.0$ N.

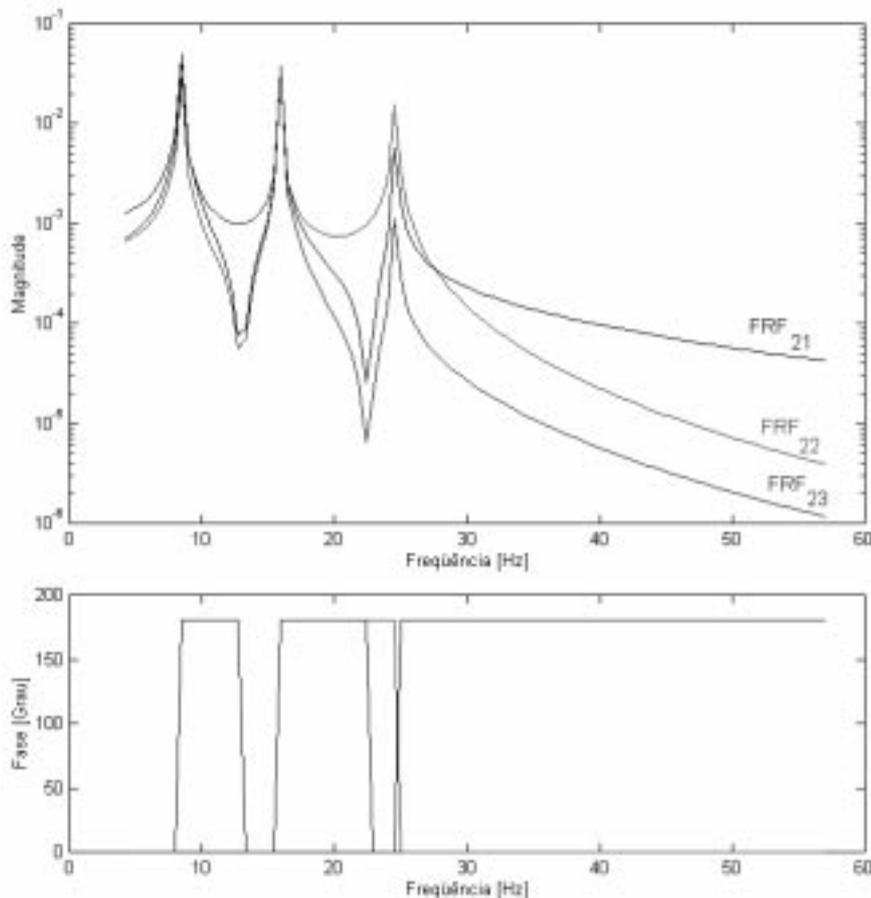


Figura (7.3) - Função resposta em frequência: sistema 3 GDL.

Nesta fase do trabalho, não se buscou a otimização da arquitetura da rede neural bem como dos seus parâmetros de treinamento. Os valores iniciais dos parâmetros de treinamento foram ajustados inicialmente como sugere Rumelhart (Rumelhart, 1986), para todos os casos testados. Posteriormente estes parâmetros foram alterados para adequar o processo de aprendizagem às particularidades de cada caso.

A rede neural back-propagation implementada foi ajustada com as seguintes características:

- arquitetura: 60 x 50 x 3
- camada de entrada e oculta: função de ativação sigmoidal
- camada de saída: função de ativação linear
- método de otimização: Método dos gradiente descendente
- taxa de aprendizado (γ): 0.001
- incremento da taxa de aprendizado = 1.05
- decremento da taxa de aprendizado = 0.60
- momentum: 0.30
- razão de erros: 1.03
- erro total admissível: 0.001

7.3 - Localização da fonte de excitação

Cada experimento é composto pelos vetores das magnitudes e das fases das respostas em frequência do sistema, respectivamente $\{M\}$ e $\{F\}$. O banco de dados destinado ao treinamento da rede neural foi composto de 30 experimentos. P é a matriz de dados de entrada para a rede neural. T é a matriz de dados de saída e representa a posição da aplicação da força de excitação. A equação (7.1) mostra a organização do banco de dados.

$$P = \begin{bmatrix} \{M\}_{m_1} & \cdots & \{M\}_{m_1} \\ \{F\}_{m_1} & \cdots & \{F\}_{m_1} \\ \{M\}_{m_2} & \cdots & \{M\}_{m_2} \\ \{F\}_{m_2} & \cdots & \{F\}_{m_2} \\ \{M\}_{m_3} & \cdots & \{M\}_{m_3} \\ \{F\}_{m_3} & \cdots & \{F\}_{m_3} \end{bmatrix}_{300 \times 30} \quad T = \begin{bmatrix} 0 & 0 & \cdots & 1 \\ 0 & 1 & \cdots & 0 \\ 1 & 0 & \cdots & 0 \end{bmatrix}_{3 \times 30} \quad (7.1)$$

A convergência da rede neural, treinada para reconhecer a posição da aplicação da força, foi atingida após 156 iterações.

Usando experimentos calculados para valores de forças que não participaram do treinamento, mas contidos entre os limites de 1.0 e 10.0 N (região do domínio de variação da

força), a rede neural reconheceu a posição da aplicação da força em 100% dos casos testados.

7.4 - Sensibilidade da Rede Neural à Presença de Ruído.

Foi montado um segundo banco de dados com a resolução em frequência da FRF reduzida para 50 pontos, ou seja $\Delta f' = 2 \Delta f$, mantendo a mesma arquitetura 60, 50 e 3 neurônios.

Adicionando ruído aleatório com diferentes níveis em ambos os bancos de dados. O nível do ruído foi ajustado com base no valor RMS de cada amostra. Os resultados obtidos estão sumarizados na tabela (7.1).

Tabela (7.1) - Localização do ponto da aplicação força: na presença de ruídos.

Arquitetura	Número de pontos da amostra	Número de experimentos	Ruído adicionado aos dados de treinamento [%]	Taxa de acerto %	Iterações até a convergência
60 x 50 x 3	100	30	0	100	156
		30	10	97	182
		30	15	78	247
		30	20	60	415
	50	30	0	90	354
		30	10	71	387
		30	15	34	478
		30	20	15	692

Pode-se observar que é elevada a insensibilidade das redes neurais à presença de ruídos em ambos os experimentos. No segundo experimento os resultados foram prejudicados devido à menor resolução em frequência. Os resultados mostram que dados com maior presença de ruído implicam em um custo de treinamento mais elevado.

A qualidade do sinal, ou seja, o quanto o sinal é bem representativo de um dado fenômeno, tem relação direta com o nível de ruído aceitável presente no sinal. Da tabela anterior observa-se que as redes neurais treinadas com dados de maior resolução (100 pontos na banda de análise) têm melhor desempenho na presença de ruído quando comparadas as redes neurais treinadas com dados menos ricos em informações (50 pontos na banda de análise). Em outras palavras, a robustez das redes neurais à presença de ruído depende da representatividade dos dados com os quais ele foi treinada.

7.5 - Influência da Compactação dos dados no Funcionamento da Rede Neural.

No capítulo 4, secção 4.9, foi mostrado o embasamento teórico de um método que permite a redução do número de variáveis de um conjunto de experimentos, reescrevendo-o em um novo espaço vetorial. O código implementado está no anexo X.

A partir do primeiro banco de dados montado, cujos experimentos contém 100 pontos, outros bancos de dados foram obtidos para diferentes taxas de redução. A mesma rede neural foi treinada com os novos bancos de dados. A tabela (7.2) mostra os resultados obtidos.

Tabela (7.2) - Localização do ponto da aplicação força: dados compactados.

Arquitetura	Taxa de compactação dos dados [%]	Fração da variância [%]	Taxa de acerto da rede [%]	Número de iterações até a convergência
60 x 50 x 3	0	100	100	156
	10	98	100	286
	20	91	100	343
	30	86	98	291
	40	83	95	385
	50	71	92	653
	60	65	87	2684
	70	57	-----	-----

A taxa de compactação está diretamente relacionada com o número de autovetores retidos pela matriz de transformação $[\bar{V}]_{r,n}$ tal como descrito pela equação (4.74) ($[\bar{V}]_{r,n} \cdot [P]_{n,p} \Leftrightarrow [\bar{P}]_{r,p}$).

Os resultados contidos na tabela (7.2) mostram que é promissor o uso desse método de compactação de dados. Para a taxa de compactação de 70% a rede não conseguiu ser treinada com os mesmos parâmetros de aprendizagem.

Na maior parte da literatura de redes neurais, o número de iterações necessário à convergência dos erros é tratado como parâmetro de qualificação de uma dada configuração da rede neural. Cabe aqui salientar que, isto só é verdade se o ponto de partida usado na busca da solução ótima for o mesmo para todas as redes que estão sendo comparadas. O grande dilema dos métodos de otimização numérica é que, pontos de partida diferentes quase sempre convergem para resultados diferentes, ainda que próximos. A rede neural implementada usa como ponto de partida, pesos das conexões gerados aleatoriamente e posteriormente normalizados; ou seja, ponto de partida diferente a cada reinicialização. Portanto, neste caso, o número de iterações não é um parâmetro válido de comparação entre os casos estudados.

Capítulo 8

Resultados e Conclusões

8.1 - Rede neural de identificação de excitação por desbalanceamento.

Fontes de excitação por desbalanceamento promovem alteração na resposta do sistema apenas na freqüência de rotação. Portanto para este tipo de excitação, as informações de magnitude e fase na freqüência de rotação, mensurados em diferentes pontos do rotor, trazem todas as informações necessárias à caracterização da fonte excitadora.

Redes neurais back-propagation com diferentes número de neurônios foram treinadas com bancos de dados compostos dos sinais de magnitude e fase mensurados na rotação de regime, por sensores instalados em diferentes posições.

A montagem do banco de dados, em todos os casos, é feita através do seguinte procedimento: cada experimento é uma coluna do banco de dados; cada coluna é composta dos sinais da magnitude e da fase dos sensores considerados. A equação (8.1) mostra como é feita a montagem do banco de dados, onde M_i é a magnitude e F_i é a fase do sinal adquirido pelo sensor i . O índice p refere-se ao número de experimentos, que neste caso vale 40.

$$[P] = \left[\begin{array}{c} \left[\begin{array}{c} M_1 \\ F_1 \\ M_2 \\ F_2 \\ \dots \\ M_i \\ F_i \end{array} \right]_1 \left[\begin{array}{c} M_1 \\ F_1 \\ M_2 \\ F_2 \\ \dots \\ M_i \\ F_i \end{array} \right]_2 \dots \left[\begin{array}{c} M_1 \\ F_1 \\ M_2 \\ F_2 \\ \dots \\ M_i \\ F_i \end{array} \right]_p \end{array} \right] \text{ e } [T] = \left[\begin{array}{c} \left\{ \begin{array}{c} disco \\ desbalc. \end{array} \right\}_1 \left\{ \begin{array}{c} disco \\ desbalc. \end{array} \right\}_2 \dots \left\{ \begin{array}{c} disco \\ desbalc. \end{array} \right\}_p \end{array} \right] \quad (8.1)$$

Os discos, como visto pela figura (6.1), de cima para baixo, foram designados os seguintes valores indicadores de suas posições (*disco*): 1, 5 e 9, respectivamente. A variável (*desbalc*) refere-se ao produto da massa de desbalanceamento pelo raio da fixação.

A definição da arquitetura da rede neural e de seus parâmetros de treinamento não segue nenhuma regra predefinida. Claramente, nestes casos a experiência do projetista com relação à natureza do problema enfrentado é de fundamental importância.

A tabela (8.1) foi gerada executando 20 experimentos que não participaram do treinamento, em algumas das redes neurais treinadas, dentre uma grande diversidade de redes, de modo a selecionar as melhores combinações de arquiteturas.

A numeração dos sensores é tal como mostrado pela figura (6.1).

Os seguintes parâmetros de treinamento foram mantidos para as várias arquiteturas treinadas:

- camada de entrada e oculta: função de ativação sigmoidal
- camada de saída: função de ativação linear
- método de otimização: Método dos gradientes descendentes
- taxa de aprendizado (γ): 0.001
- incremento da taxa de aprendizado = 1.05
- decremento da taxa de aprendizado = 0.60
- momentum: 0.30
- razão de erros: 1.03
- erro total admissível: 0.001

A tabela anterior foi gerada executando as diferentes arquiteturas de redes neurais treinadas com diferentes bancos de dados.

A resposta da rede neural é o valor do desbalanceamento e a indicação da sua localização.

A coluna da média do erro quadrático (E_p) é obtida acumulando o quadrado do erro cometido pela rede, após a execução dos dados reservados para sua qualificação, segundo a equação (8.2), onde n é o número de experimentos, e E_i^2 é o quadrado do erro de cada experimento.

$$E_p \equiv 100 \frac{1}{n} \sum_{i=1}^n E_i^2 \quad (8.2)$$

No caso presente, E_p é obtido apenas com a média do erro quadrático dos neurônios da camada de saída da rede neural, correspondentes aos parâmetros da força de desbalanceamento. O erro cometido pelo neurônio de indicação da localização do desbalanceamento não participa do cálculo de E_p . Então E_p representa neste caso a incerteza do valor do desbalanceamento.

Observa-se que para determinação da localização e do valor do desbalanceamento, os sinais adquiridos apenas de dois planos são suficientes. As redes neurais treinadas com sinais mensurados em mais do que dois planos, apresentam arquitetura com maior número de neurônios e são menos eficientes, quando comparadas às redes neurais treinadas com dois planos de medição.

Uma característica dessa classe de problema é que, a localização do desbalanceamento é mais facilmente determinada do que o seu valor.

A arquitetura 4x12x2, treinada com os sinais dos sensores número 1 e 3, apresenta melhor performance global que as demais.

Tabela (8.1) - Qualificação das redes neurais de identificação de desbalanceamento.

Sensores	Arquitetura	Ep % (amplitude)	Localização: Taxa média de acerto %
[1,3]	4x4x2	8.18	100
	4x8x2	6.56	100
	4x12x2	3.05	100
	4x16x2	7.65	100
	4x20x2	11.87	100
[4, 5]	4x4x2	9.11	100
	4x8x2	6.96	100
	4x12x2	3.21	100
	4x16x2	8.51	100
	4x20x2	12.37	100
[1,2,3]	4x36x2	16.52	94
	6x42x2	19.76	97
	6x48x2	18.53	94
	6x54x2	25.02	91
[1,2,3,4,5]	10x10x2	123.65	87
	10x20x2	57.93	85
	10x30x2	62.81	71
	10x40x2	134.76	87

8.2 - Rede neural de identificação de excitação assíncrona no mancais.

O sinal da resposta em frequência de um sistema na presença de excitação forçada, apresenta um pico de maior concentração de energia na posição de frequência correspondente à frequência da força excitadora. Com base nessa característica, do mesmo modo que na excitação por desbalanceamento, toda as informações necessárias à caracterização da força excitadora, estão contidos nos valores de magnitude e fase correspondente à frequência de excitação.

No caso de excitação assíncrona, algumas situações podem dificultar a implementação de um algoritmo de seleção dos valores de magnitude e fase dos picos contidos no sinal da resposta em frequência do sistema. Considere a situação em que a frequência da força excitadora esteja muito próxima da frequência de rotação ou de uma de suas harmônicas. Neste caso, se a distancia em frequência que separa os picos da força excitadora e da rotação, for pequena, fica

difícil separar as duas freqüências, devido à natureza discreta dos dados experimentais.

No caso presente, optou-se por usar uma versão compactada da função resposta em freqüência. O algoritmo utilizado foi tal como descrito na secção 4.9. Vários níveis de redução dos dados foram estudados em conjunto com diferentes arquiteturas de redes neurais. A montagem dos bancos de dados foi tal como mostrado pela equação (8.3).

$$[P] = \left[\begin{array}{c} \left[\begin{array}{c} M_1 \\ F_1 \\ M_2 \\ F_2 \\ \dots \\ M_i \\ F_i \end{array} \right]_1 \\ \left[\begin{array}{c} M_1 \\ F_1 \\ M_2 \\ F_2 \\ \dots \\ M_i \\ F_i \end{array} \right]_2 \\ \dots \\ \left[\begin{array}{c} M_1 \\ F_1 \\ M_2 \\ F_2 \\ \dots \\ M_i \\ F_i \end{array} \right]_p \end{array} \right] \text{ e } [T] = \left[\begin{array}{c} \left[\begin{array}{c} mancal \\ amplit \\ freq \end{array} \right]_1 \\ \left[\begin{array}{c} mancal \\ amplit \\ freq \end{array} \right]_2 \\ \dots \\ \left[\begin{array}{c} mancal \\ amplit \\ freq \end{array} \right]_p \end{array} \right] \quad (8.3)$$

Onde $[P]$ e $[T]$ são respectivamente as matrizes de entrada e saída necessárias ao treinamento da rede, montadas com 40 experimentos. M_i e F_i são os valores de magnitude e fase da FRF medidas pelo sensor i . As variáveis (*amplit*) e (*freq*) são os valores de amplitude e freqüência da força excitadora. A variável (*mancal*) a designa a posição de atuação da força excitadora, mancal superior (*mancal* = 1), mancal inferior (*mancal* = 5).

A tabela (8.2) mostra algumas das redes (sigmoide - sigmoide - linear), treinadas com diferentes bancos de dados gerados e testadas com 20 experimentos que não participaram do treinamento. Neste caso, Ep significa a incerteza dos valores de amplitude e da freqüência da força de excitação.

Os parâmetros necessários ao aprendizado das redes neurais foram mantidos constantes, ajustados da seguinte forma:

- método de otimização: Método dos gradientes descendentes
- taxa de aprendizado (γ): 0.001
- incremento da taxa de aprendizado = 1.01
- decremento da taxa de aprendizado = 0.60
- momentum: 0.70
- razão de erros: 1.03
- erro total admissível: 0.001

As redes neurais treinadas com os sinais provenientes da combinação dos sensores [1,3] e [4,5] mostraram-se mais eficientes que os demais; o que indica que tais sensores estão posicionados em locais sensíveis à excitação.

Tabela (8.2) - Qualificação das redes neurais de caracterização de excitação assíncrona.

Sensores utilizados	Taxa de redução dos dados [%]	Razão da variância [%]	Arquitetura da rede neural	Ep % (amplitude e frequência)
[1,3]	0	100	10x10x3	----
			10x20x3	52.1
			10x30x3	44.5
	10	97	10x10x3	42.8
			10x20x3	57.9
			10x30x3	31.3
	20	95	10x10x3	34.11
			10x20x3	14.26
			10x30x3	36.47
	30	91	10x10x3	8.64
			10x20x3	5.31
			10x30x3	4.48
	40	88	10x10x3	21.47
			10x20x3	30.15
			10x30x3	53.4
	50	71	10x10x3	64.31
			10x20x3	87.17
			10x30x3	98.18

Sensores utilizados	Taxa de redução dos dados [%]	Razão da variância [%]	Arquitetura da rede neural	Ep % (amplitude e frequência)
[1,2,3]	0	100	20x40x3	----
	10	99		67.3
	20	95		56.8
	30	91		38.2
	40	87		71.4
	50	78		61.03
[4,5]	0	100	10x20x3	----
	10	98		34.3
	20	95		15.2
	30	89		9.2
	40	78		9.4
	50	64		17.6
[1,2,3,4,5]	0	100	40x50x3	85.3
	10	99		76.9
	20	97		72.8
	30	91		91.8
	40	87		---
	50	72		---

Os bancos de dados compostos por sinais de vários sensores, levaram ao treinamento de redes neurais com maior número de neurônios que as demais, sem contudo apresentar ganho de eficiência nos seus resultados. Este fato tem pelo menos duas razões. A primeira razão diz respeito à sensibilidade dos sensores localizados mais ao centro do rotor, que de um modo geral, captam com menor intensidade a resposta do sistema devido à excitações localizadas nos mancais. A segunda razão, é de natureza numérica; um grande número de variáveis que não acrescentam informações ao banco de dados, levam ao treinamento redes neurais com maior número de neurônios, portanto passíveis de apresentarem o fenômeno do sobreajuste, tal como mostrado na figura (4.8).

A arquitetura 10x20x3, treinada com sinais dos sensores 1 e 3, com taxa de compactação dos dados de 30%, apresentou a melhor performance ($Ep = 5.31$) que a mesma arquitetura, treinada para a mesma taxa de compactação, porém com os sinais dos sensores 4 e 5, ($Ep = 9.29$), ainda que estes sensores estejam localizados nos pontos de excitações. Uma das razões que explicam esse resultado, é fato de que nos mancais existem grandes concentrações de massa, logo uma excitação localizada num dos mancais, é medida com pouca intensidade no

outro mancal.

A arquitetura 10x30x3, treinada com sinais dos sensores 1 e 3, com taxa de compactação dos dados de 30%, apresentou a melhor performance dentre todas as combinações testadas, $Ep = 4.48$.

8.3 - Rede neural de identificação de excitações por forças magnéticas.

O sinal da FRF (função resposta em frequência) característico da excitação por forças magnéticas contém vários picos de frequência com substancial nível de energia, além das frequência fundamental de excitação. Este fato é devido a natureza não linear do campo magnético, produzindo uma força de natureza periódica. A compactação de dados aqui utilizada foi efetuada como descrita no item 4.9.

Foi tentado, porém não foi possível, a caracterização da posição da força excitadora uma vez que o dispositivo de excitação magnética usado, um ímã permanente, não tem a capacidade de excitar o rotor em pontos de grande concentração de massa tais como mancais e no disco 3. Em virtude de tal fato, assumiu-se que este tipo de excitação ocorre apenas no disco intermediário do rotor, na direção oposta ao sensor de número 2.

Esta rede neural deverá fornecer a frequência principal de excitação e a sua amplitude pico a pico. Diferentes arquiteturas de rede neurais back-propagation, com funções de ativação sigmoide - sigmoide - linear, cujos parâmetros de treinamento foram ajustados da seguinte forma:

- método de otimização: Método dos gradientes descendentes
- taxa de aprendizado (γ): 0.001
- incremento da taxa de aprendizado = 1.05
- decremento da taxa de aprendizado = 0.60
- momentum: 0.80
- razão de erros: 1.02
- erro total admissível: 0.001

$$[P] = \left[\begin{array}{c} \left[\begin{array}{c} M_1 \\ F_1 \\ M_2 \\ F_2 \\ \dots \\ M_i \\ F_i \end{array} \right]_1 \\ \left[\begin{array}{c} M_1 \\ F_1 \\ M_2 \\ F_2 \\ \dots \\ M_i \\ F_i \end{array} \right]_2 \\ \dots \\ \left[\begin{array}{c} M_1 \\ F_1 \\ M_2 \\ F_2 \\ \dots \\ M_i \\ F_i \end{array} \right]_p \end{array} \right] \text{ e } [T] = \left[\begin{array}{c} \left\{ \begin{array}{c} \text{amplit} \\ \text{freq} \end{array} \right\}_1 \\ \left\{ \begin{array}{c} \text{amplit} \\ \text{freq} \end{array} \right\}_2 \\ \dots \\ \left\{ \begin{array}{c} \text{amplit} \\ \text{freq} \end{array} \right\}_p \end{array} \right] \quad (8.4)$$

Diferentes bancos de dados foram montados variando a posições dos sensores e a taxa

de compactação dos dados, tal como mostrado pela equação (8.4). $[P]$ e $[T]$ são as matrizes de entrada e de saída necessárias ao treinamento da rede, montadas com 40 experimentos. ($amplit$) e ($freq$) são respectivamente a amplitude e a frequência da força de excitação.

A Tabela (8.3) mostra o resultado da execução de 20 novos experimentos, em algumas das redes neurais que conseguiram êxito no treinamento.

Da tabela (8.3), observa-se que o sinal unicamente do sensor 2 foi capaz de levar à identificação da amplitude e frequência da força excitadora, uma vez que os pontos de excitação e de medição são coincidentes. As redes neurais treinadas com bancos de dados montados com os sinais dos sensores 1 e 3, globalmente obtiveram melhor performance que as demais, já que estão mais próximos do ponto de excitação do que os sensores 4 e 5.

Os sensores 4 e 5, como mostrado pela figura (7.1), estão situados nos mancais, dotados de massas expressivas, logo, menos sensíveis a uma perturbação de pequena magnitude que atua num ponto intermediário do rotor.

Tabela (8.3) - Qualificação das redes neurais de identificação de excitação magnética.

Sensores	Taxa de redução dos dados [%]	Razão da variância [%]	Arquitetura da rede neural	Ep [%] (amplitude e frequência)	Sensores	Taxa de redução dos dados [%]	Razão da variância [%]	Arquitetura da rede neural	Ep % (amplitude e frequência)
[2]	10	95	27x54x2	7.81	[1,2,3]	0	100	70x140x2	2.54
			27x108x2	30.53		10	97	63x252x2	8.89
[1,3]	0	100	50x100x2	64.62		20	94	56x168x2	8.37
			50x150x2	27.63		30	92	49x98x2	48.79
			50x200x2	47.10		40	90	42x126x2	3.99
	10	98	45x45x2	18.56		50	78	35x175x2	13.66
			45x90x2	2.66	[4,5]	0	100	50x200x2	30.33
			45x135x2	20.81		10	98	45x180x2	28.62
	20	96	40x40x2	34.11		20	94	40x160x2	27.00
			40x80x2	4.26		30	89	35x140x2	22.24
			40x120x2	36.47		40	86	30x150x2	21.82
	30	92	35x105x2	58.49		50	79	25x100x2	41.22
			35x140x2	49.21	[1,2,3,4,5]	0	100	110x440x2	17.07
			35x175x2	2.01		10	98	99x396x2	17.31
40	86	30x60x2	16.09	20		95	88x440x2	18.84	
		30x150x2	40.58	30		93	77x308x2	10.92	
		30x30x2	50.76	40		90	66x330x2	6.78	
50	71	25x50x2	29.59	50		84	55x220x2	13.42	
		25x75x2	14.17						
		25x100x2	48.58						

A grande deficiência das redes neurais back-propagation está relacionado ao seu aprendizado. Para um mesmo banco de dados, arquitetura da rede e o ajuste fino dos parâmetros de treinamento podem significar a convergência rápida, lenta ou mesmo a não convergência dos algoritmos de otimização. Porém, o aspecto de fundamental importância, quando se trabalha com rede neural, é a escolha e a manipulação adequada dos dados fornecidos à ela.

Dados adequados implicam em arquitetura menor da rede neural, menor esforço computacional durante o seu treinamento e menor erro de execução.

8.4 - Rede neural de classificação.

Esta rede neural tem por finalidade a indicação da classe a qual pertence um dado sinal a ela apresentado. Esta rede neural foi treinada para identificar quatro classes distintas, são elas, classe 1: sinal característico da bancada sem a presença de defeitos, o que corresponde ao estado normal de operação do rotor; classe 2: sinal característico de excitação assíncrona nos mancais, classe 3: sinal característico de excitação por forças magnéticas, e finalmente a classe 4: sinal característico de excitação por desbalanceamento.

Como nos casos das redes neurais anteriormente estudadas, os sinais dos sensores 1 e 3 mostraram-se sensíveis aos diversos tipos de excitação aplicados; logo estão naturalmente eleitos para fornecerem dados à rede de classificação. Para a metodologia aqui aplicada, os demais sensores são desnecessários.

Foram adquiridos 30 amostras na rotação de regime com o rotor não intencionalmente perturbado. Dentre estas, 15 amostras foram retidas para avaliação da rede neural.

O treinamento dessa rede neural foi executado usando um banco de dados formado por 15 amostras representativas de cada uma das classes, totalizando 60 amostras.

Níveis diferentes de redução do banco de dados, tal como discutido na seção 4.9, foram usados no treinamento de diferentes arquiteturas de rede neural Back-propagation, de três camadas, com funções de ativação: sigmoideal, sigmoideal e linear. Os parâmetros de treinamento foram os seguintes:

- método de otimização: Método dos gradientes descendentes
- taxa de aprendizado (γ): 0.001
- incremento da taxa de aprendizado = 1.01
- decremento da taxa de aprendizado = 0.60
- momentum: 0.70
- razão de erros: 1.03
- erro total admissível: 0.001

A montagem dos bancos de dados foi executada segundo a equação (8.5); onde $[P]$ é matriz dos dados de entrada para a rede, composto pelos sinais dos sensores 1 e 3, e $[T]$ é a matriz dos dados de saída. São nulos todos de cada coluna da matriz $[T]$, a menos do elemento correspondente ao neurônio indicador da classe, cujo valor é unitário.

$$[P] = \left[\begin{array}{c|c|c|c} \begin{matrix} M_1 \\ F_1 \\ M_3 \\ F_3 \end{matrix} \\ \begin{matrix} M_1 \\ F_1 \\ M_3 \\ F_3 \end{matrix} \\ \dots \\ \begin{matrix} M_1 \\ F_1 \\ M_3 \\ F_3 \end{matrix} \end{array} \right]_1 \dots \left[\begin{matrix} M_1 \\ F_1 \\ M_3 \\ F_3 \end{matrix} \right]_p \quad \text{e} \quad [T] = \left[\begin{array}{c|c|c|c} \begin{matrix} 1 \\ 0 \\ 0 \\ 0 \end{matrix} \\ \begin{matrix} 0 \\ 1 \\ 0 \\ 0 \end{matrix} \\ \dots \\ \begin{matrix} 0 \\ 0 \\ 0 \\ 1 \end{matrix} \end{array} \right]_1 \dots \left[\begin{matrix} 0 \\ 0 \\ 0 \\ 1 \end{matrix} \right]_p \quad (8.5)$$

A Tabela (8.1) mostra o resultado da execução das redes treinadas, para um conjunto de 75 experimentos que não participaram do treinamento. As células vazias correspondem aos casos de insucesso no treinamento. Outras combinações de funções de ativação foram testados, porém, não apresentaram melhores resultados. Como destacado na tabela (8.4), a rede neural que obteve maior taxa de acerto (70%) foi a de arquitetura 40x80x4, usando banco de dados sem redução.

Tabela (8.4) - Rede neural de classificação.

Taxa de redução [%]	Razão da variância [%]	Arquitetura	Taxa de acerto [%]	Arquitetura	Taxa de acerto [%]
0	100	40x40x4	36	40x120x4	20
		40x80x4	70	40x160x4	-----
10	83	36x36x4	42	36x108x4	7
		36x72x4	47	36x144x4	13
20	72	32x32x4	9	32x96x4	26
		32x64x4	21	32x128x4	-----
30	63	28x28x4	16	28x84x4	2
		28x56x4	2	28x112x4	-----

A taxa da compactação de dados, através do algoritmo estatístico, afeta de modo significativo o aprendizado das redes neurais classificadoras. Este efeito deve-se à baixa correlação entre os dados que compõem o banco de dados; níveis ainda pequenos de redução leva ao descarte de informações importantes. De fato, o aprendizado de uma rede classificadora é realizado com um banco de dados formado por um conjunto de experimentos de classes diferentes. Em virtude desse fato, optou-se então por modificar a forma de montagem do banco de dados, de modo a reduzir o número de variáveis.

Os sinais de FRF são obtidos por sensores de deslocamento. Nas frequências naturais a fase entre uma força e o deslocamento na ressonância é 90 graus. Se multiplicarmos os valores de magnitude pelos respectivos valores de fase, estaríamos minimizando a influencia da oscilação da fase entre 0 e 180 graus, nos pontos de pequena amplitude da força, ou seja, fora das ressonâncias. A nova matriz $[\bar{P}]$ tem então a metade do número de variáveis de $[P]$.

$$[\bar{P}] = \left[\left\{ \begin{matrix} M_1 \cdot F_1 \\ M_3 \cdot F_3 \end{matrix} \right\}_1 \left\{ \begin{matrix} M_1 \cdot F_1 \\ M_3 \cdot F_3 \end{matrix} \right\}_2 \dots \left\{ \begin{matrix} M_1 \cdot F_1 \\ M_3 \cdot F_3 \end{matrix} \right\}_p \right] \text{ e } [T] = \left[\left\{ \begin{matrix} 1 \\ 0 \\ 0 \\ 0 \end{matrix} \right\}_1 \left\{ \begin{matrix} 0 \\ 1 \\ 0 \\ 0 \end{matrix} \right\}_2 \dots \left\{ \begin{matrix} 0 \\ 0 \\ 0 \\ 1 \end{matrix} \right\}_p \right] \tag{8.6}$$

Uma novo conjunto de arquiteturas foram treinadas para a matriz $[\bar{P}]$. O resultado da execução de 75 experimentos que não participaram do treinamento, em algumas das redes treinadas, estão resumidos na tabela (8.5). Os parâmetros de treinamento foram mantidos os mesmos do caso anterior.

Tabela (8.5) - Rede neural de classificação: dados não compactados.

Arquitetura	Taxa de acerto %	Arquitetura	Taxa de acerto %
10x10x4	49	20x40x4	95
10x20x4	55	20x80x4	75
10x30x4	59	30x30x4	62
10x40x4	66	30x60x4	65
20x20x4	88	30x90x4	54

A arquitetura 20x40x4 acertou 95% dos 75 experimentos a ela apresentados, portanto o novo arranjo do dados mostrou-se eficiente, permitindo inclusive o treinamento de redes neurais com menor número de neurônios.

8.5 -Conclusões.

Considerando as particularidades de cada uma das redes treinadas, o fluxograma geral interligando a rede de classificação com as redes de diagnóstico, é mostrado pela figura (8.1).

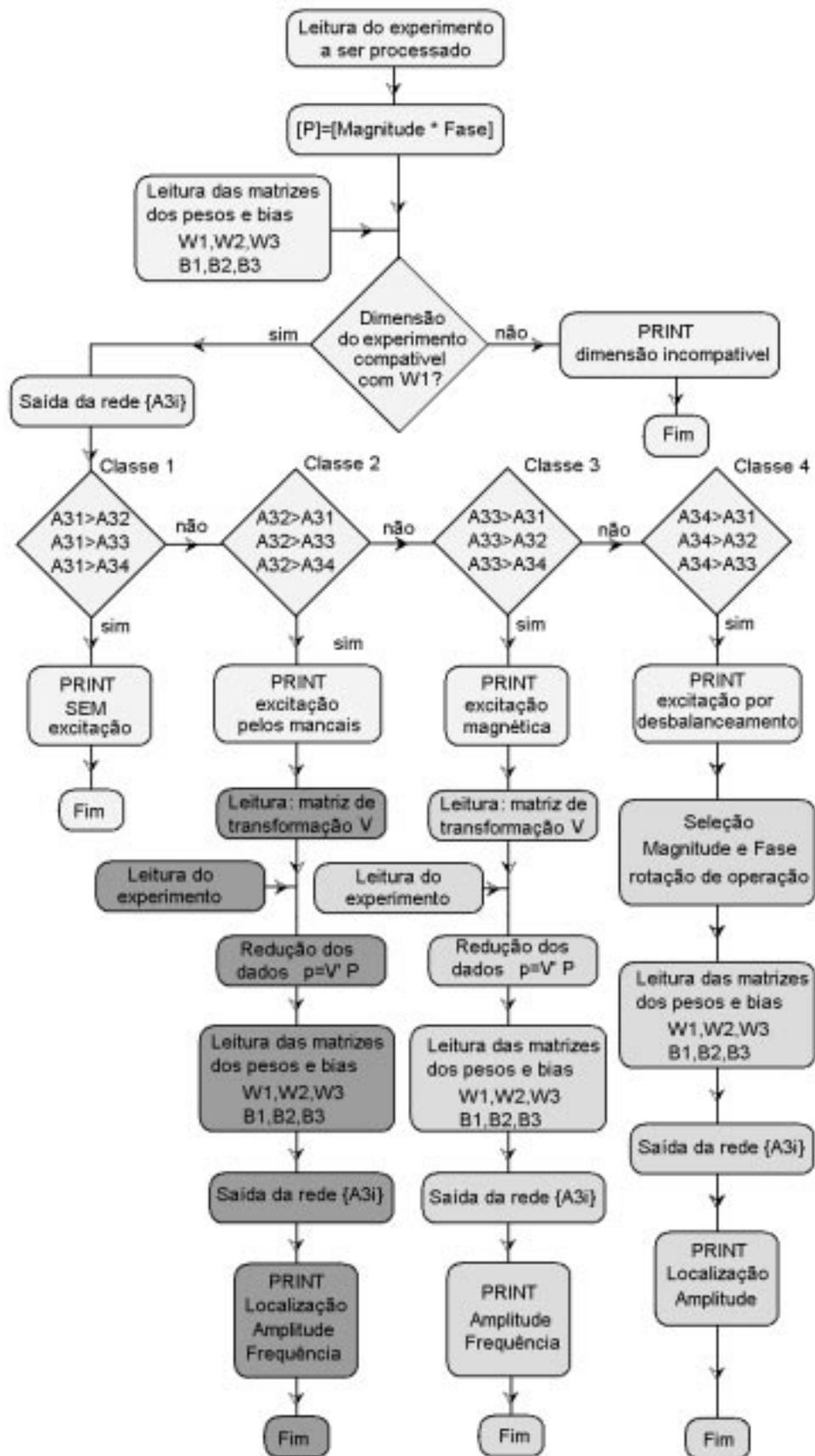


Figura (8.1) - Fluxograma geral do processo automatizado de diagnostico de defeitos.

O capítulo 4 procurou mostrar o embasamento matemático das redes neurais artificiais. Por ser uma técnica ainda recente, cujos estudos teóricos ainda estão em desenvolvimento, há muitos aspectos que precisam ser melhorados. Neste trabalho foi dada ênfase às redes neurais com treinamento supervisionado pelo algoritmo back-propagation. O enfoque foi o da aplicação, sem a pretensão de desenvolver novas técnicas de treinamento ou novos tipos de redes neurais.

Como mostrado pelas tabelas anteriores, diferentes topologias de redes neurais foram treinadas com diferentes bancos de dados. O objetivo principal desses testes foi o de selecionar as topologias mais adequadas à natureza do fenômeno tratado.

A velocidade de treinamento foi preocupação secundária. Os parâmetros de treinamento tais como, taxa de aprendizagem, coeficiente de momentum, etc., foram ajustados experimentalmente.

O bom condicionamento numérico dos sinais mensurados pode determinar o êxito do treinamento ou não de uma rede neural. Seja por exemplo os seguintes dados: magnitude igual a 0.0034 e fase igual a 97 graus, melhor seria magnitude igual a 0.0034 e fase igual a 1.6929693 radianos.

Das tabelas (8.2) e (8.3) é possível observar que com sinais de vibração medidos pelos sensores 1 e 3, conseguem treinar redes neurais com performance razoável, ainda que submetidos à níveis mais elevados de redução. Está é uma clara indicação de que nestes sinais, as informações das excitações estão fortemente evidenciadas, o que permite que redes neurais com um número menor de neurônios sejam treinadas com sucesso.

A quantidade e a qualidade dos dados necessários ao treinamento das redes neurais é um item de fundamental importância. A obtenção dos dados experimentais, foi assistida de todas as preocupações sob o aspecto técnico, para garantir a qualidade do sinal.

Procurou-se gerar bancos de dados com sinais representativos da mais ampla faixa de incidência de cada uma das excitações estudadas, já que as redes neurais, por definição, interpolam resultados. A perfeita caracterização das diversas classes defeitos é necessária para a composição do banco de dados.

A utilização dos sinais no domínio da frequência foi escolhido por ser de fácil obtenção e seus parâmetros de magnitude e fase são sensíveis aos tipos de excitação estudados.

A metodologia utilizada mostrou-se eficaz no processo de diagnóstico e caracterização das excitações aplicadas.

Do ponto de vista qualitativo, as redes neurais conseguem classificar e indicar a localização de uma dada excitação, com grande êxito. Na quantificação dos valores de frequência e amplitude da excitação, as redes neurais apresentam erros como qualquer ferramenta numérica de interpolação.

A presença simultânea de duas ou mais fontes de excitação é desconsiderada pela rede neural que, indica uma fonte inexistente ou indica apenas a mais pronunciada das excitações.

Durante a execução, as redes neurais identificam com sucesso, para os casos testados, os parâmetros contidos em sinais considerados ruidosos até o limite de 20 % do nível RMS do sinal original.

As redes neurais artificiais não representam uma ameaça ao desenvolvimento de outras técnicas de diagnóstico e caracterização de excitações de sistemas, tais como transformada de wavelet e outras. Pelo contrário, o uso das redes neurais é um incentivo ao conhecimento da fenomenologia do problema tratado, uma vez novos e mais significativos parâmetros são disponibilizados possibilitando às redes neurais, aprendizagem mais rápidas e mais precisas.

Como sugestões de prosseguimento desse trabalho, alguns itens podem ser destacados:

- Compactação dos dados.

Uma forma alternativa que apresenta alta taxa de compactação de dados para sinais no domínio do tempo e da frequência pode ser conseguido utilizando os coeficientes das funções wavelets ajustadas para representar o sinal.

- Análise de sensibilidade.

O uso de técnicas de análise de sensibilidade para identificar os pontos de medições mais adequados para a medição da resposta do sistema aos diferentes tipos de excitação, pode introduzir um ganho significativo na qualidade dos dados adquiridos, economizando ciclos de treinamento e propiciando o treinamento de redes neurais mais robustas à presença de ruídos. Neste contexto, o uso do modelo computacional para a simulação das respostas aos diferentes tipos de excitações é uma ferramenta importante.

- Redes neurais.

Com relação às redes neurais back-propagation, a utilização de métodos de busca aleatória de mínimos menos vulnerável à presença de mínimos locais, tal como o método de simulated annealing.

Com relação a rede de classificação, a evolução natural será a utilização de redes auto-organizadas, como descrita no item 4.7. Tais redes são essencialmente ferramentas de classificação de padrões.

Referências Bibliográficas

“ANSYS™ User Manual”, Revision 5.2, December 1992.

“GENESIS™ User Manual”, VMA Engineering.

Alguindigue, A., Loskiewicz-Buczak, A., and Uhrig, R.E., “Monitoring and Diagnosis of Rolling Element Bearing Using Artificial Neural Networks”, IEEE Transaction on Industrial Electronics, Vol. 40, No. 2, April 1993.

Alguindigue, I.E. e Uhrig, E.R. “Compression of Spectral Signatures Using Recirculation Networks”, Scientific Computing and Automation, Vol. 7, No. 7, 1991

Anderson, T. W. “An Introduction to Multivariate Statistical Analysis”. John Wiley & Sons, INC. New York. 1958.

Araújo, C. A.; “Controle da Vibração de Rotores Flexíveis Usando Forças Magnéticas Externas”, tese de mestrado em Engenharia Mecânica, UFU, Março de 1993.

Assis, E. G. e Steffen, V. , “Parameter Identification in Rotor Bering Systems”, IMAC - 17h International Modal Analysis Conference. Society for Experimental Mechanics. 1999.

B. Windrow e M. E. Hoff, “Adaptive switching circuits”, 1960 IRE WESCON Convention Record, New York, 1960, 96-104.

Ball, K. W, Tomlinson, G. R., “Neural Networks for Fault Location”, Mechanical Systems and Signal Processing, 1992.

Bathe, K. and Wilson, E. L., “Numerical Methods in Finite Element Analysis”. 1986.

Bendat, Julius S., Piersol, Allan G. “Random Data: Analysis and Measurement Procedures”. 2o Edition, John Wiley & Sons., USA. 1986.

Bernieri, A., M. D’Apuzzo; L. Sansone, L. e Savastano, M. , “A Neural Network Approach for Identification and Fault Diagnosis on Dynamic Systems”, IEEE Transactions on Instrumentations and Measurement, Vol. 43, No. 6..December 1994.

Berriet, R. e Link, M. “Updating of Modal Damping, Physical Stiffness and Mass Parameters from Measured Frequency Response Functions”, ISMA19 - Tools for Noise and Vibration Analysis, pag. 291 -304

Berthier, Ferraris e Lalanne; "Prediction of Critical Speeds, Unbalance and Nonsynchronous Forced Response of Rotors", The Shock and Vibr. Bull., part 4, 103-111, May 1983.

Berthier, P. ; Ferraris, G.; Der Hagopian, J.;Lalanne, M., "Application d'une methode pseudo-modale. La determination du comportement dynamique des monorotors et des multrotors coaxiaux". Proceeding 7th Word IFToMM Congress, Pergamon Press, 1987.

Berthier, P.;Ferraris, G.; Lalanne, M.;"Application d'une m,thode pseudo-modale . la determination du comportement dynamique des monorotors et des multrotors coaxiaux". Proceeding 7th Word IFToMM Congress, Pergamon Press, 1987.

Cawley, P. e Adams, R.D. "The Location of Defects in Structures from Measurements of Natural Frequencies", Journal of Strain Analysis, 14(2), pp. 49-57,1979.

Craig, Jr., Roy, R., "Structural Dynamics. An Introduction to Computer Methods", John Wiley & Sons, New York, N. Y. 1981.

Davidon, W. C.; Variable Method for Minimization, Research and Development Report ANL-5990, 1969.

Eengels, R. C., "Finite Element Modeling of Dynamic Behavior of Some Basic Structural Members", Journal of Vibration and Acoustics, pp 236-241, Vol. 114, Jan.1991.

Farge, M.."Wavelet Transform and their Aplication to Turbulence",Annual Review of Fluid Mechanics. Vol. 24, pp 395-457. 1992.

Faria, Mário Luiz M.; "Uma Contribuição aos Procedimentos de Otimização Aplicados a Sistemas Mecânicos" , dissertação de mestrado, U.FU, 1991.

Felman, M.."Non-linear Sistem Vibration Analisis Using Hilbert Transform: Free Vibration Analisis Method 'FREEVIB'".Mechanical Systems and Signal Processing,. Vol. 8, pp. 119-127. 1994.

Fletcher, R., Powell, M. J. D., "A Rapidly Convergent Descent Method for Minimization", Computer Journal, 6 (2) : 163-168, 1963.

Fregolent, A., Sestieri, A., "Force Identification from Vibration Measurements in the Wavenumber Domain", ISMA21, International Conference in Noise and Vibration Engineering. Leuven, Belgium. 1996 .

Gonzalez, P. S., "Diagnostico de Fallas en Maquinas Rotatorias: Estado del Arte",VI Congresso Nacional de Ingenieeria Mecanica, Chile. 1996.

Gosiewski, Z.; "Automatic balancing of Flexible Rotors, Part I: Teorical Background", Journal of Sound and Vibration, 100 (4), p. 551 - 567, 1985.

Green, G. G. R. e Manchada, S., "System Identification and Minimal Model Realization Using Recurrent Neural Networks", Identification in Engineering Systems Proceedings of the Conference at Swansea, March 1996.

Grossberg, S., "Neural pattern discrimination", J. Theory Biol. vol. 27, pp. 291-337. 1970.

Haykin, S.. "Neural Networks A Comprehensive Foundation". MacMillian Publishing Company. New York. 1994.

Hewlett Packard Data Aquisition, HP 3563R Programer's Toolkit, Programming in MIL - Measurement Interface Library. Hewlett-Packard Company. December 1993. Washington, USA.

Hopfield, J. J. e Tank, D. W., "'Neural' computation of decisions in optimization problems", Biological Cybernetics, Vol. 52, pp. 142-169. 1985.

Hopfield, J. J., "Neural networks and physical system with emergent collective computational abilities", Proceedings of the National Academy of Sciences, Vol. 79 , pp. 2554-2558, 1982.

Hoskins, J. C. , Kaliyur, K. M. and Himmelblau,D. M., "Fault Diagnosis in Complex Chemical Plants Using Artificial Neural Networks", AIChE Journal, Vol. 37, No. 1, January 1994.

Jaswinski, A.H., Stochastic Process and Filtering Theory. Academic Press, New York. 1970.

Kalman, R. E., "A new Approach to Linear and Prediction Problems", Trans. ASME, Series D, Journal of Basic Engineering, Vol. 82, pp. 35-45. 1960.

Kröse, B. J. A. e Vander Smagt, P. P., "An Introduction to Neural-Networks"; The University of Amsterdam, 1993.

Lacerda, H. B. e Steffen, V. Jr.; Balanceamento de Rotores Flexíveis Usando Otimização, U.F. Uberlândia ,Junho de 1991.

Lacerda, H. B.; "Racionalização do Método dos Coeficientes de Influência aplicado ao Balanceamento de Rotores Flexíveis" , dissertação de mestrado, U.F. Uberlândia , 1991.

Lalanne, M. e Ferarris,G., "Rotordynamics Prediction in Engineering", John Wiley & Sons, New York, 1990.

Lalanne, M., Berthier, P. and Hagopian, J. D., "Mechanical Vibrations for Engineers", John Wiley and Sons, New York, 1984.

Lalanne, Queau e Tran; "Dynamique des Rotors: Vitesses Critiques , Instabilités"; Laboratoire de Mécanique des Structures - I.N.S.A. (França), pp.105-113. 1982.

Lallement, Lecoanet e Steffen; "Vibration de Rotors sur Paliers à Matrice de Raideur non Symétrique", Mech and Mach Theory, vol. 17(1), pp. 4755, 1982.

Lee, S.; e Kill, R, "Bidirectional continuous associator based on Gaussian potential function network", in Proc. IEEE / INNS Int. Joint Conf. Neural Networks, vol. I, pp. 45-54. 1989.

Lembregts, F., Top, J. e Neyrinck, F., "Adaptative Resembling for Off-line Signal Processing", ISMA21, pp. 1171-1182. International Conference in Noise and Vibration Engineering. Leuven, Belgium. 1996.

Lépore N., F. P. and Steffen, V. Jr.; "A Control Force to Reduce Magnetically Excited Rotor Vibration", Mécanique Matériaux Électricité, Vol. 4, pp. 90 - 93, April 1984.

Liu, T. I. e Mengel, J. M.; "Inteligent Monitoring of Ball Bearing Conditions". Mechanical System and Signal Processing. Vol. 6, pp. 419 - 431. 1992.

Liu, T. I.; Singonahalli, J. H. e Iyer, N. R., "Detection of Roll Bearing Defects Using Expert", Mechanical Systems and Signal Processing. Vol. 10(5), pp. 595 - 614. 1996.

Lucifredi, A. e Hou, Z. "The Construction of Neural Network for Bilinear Systems Identification and a Possible Use to Crack Detection", ISMA19 - Tools for Noise and Vibration Analysis, pp. 1139 -1149. 1996.

Master, Timothy, "Advanced Algorithms for Neural Network: A C++ Sourcebook". John Wiley & Sons, INC. New York. 1995.

McCulloch, W. S. e Pitts, W., "A logical calculus of the ideas immanent in nervous activity", Bulletin of Mathematical Biophysics. Vol. 5 , pp. 115-119. 1943.

Minh, T. D., "Etude du Comportement Dynamique des Rotors Flexibles", These pour obtenir grade de Docteur de L'Universite de Claude Bernard Lyon I - L'ecole Centrale de Lyon, Lyon-France.1981.

Minsky, M. e Papert, S., "Perceptrons: An Introduction to Computational Geometry", MIT Press, 1969.

Mistry, S. I. e Nair. S.S., "Neural Network Designs for Partially Known Dynamic Systems", Journal of Engineering for Industry. Vol. 116/407, August 1994.

Nauerz, A. e Fritzen, C. P., " Error Localization Using Power Spectral Densities", ISMA21, International Conference in Noise and Vibration Engineering. Leuven, Belgium, 1996.

Neto, Atair Rios, "Stochastic Optimal Linear Parameter Estimation and Neural Network Training in Systems Modeling", Sociedade Brasileira de Ciências mecânicas, Vol. XIX - No. 2. pp. 138 - 146, 1997.

Nussbaumer, H. J., "Fast Fourier Transform and Convolution Algorithms". New York: Springer-Verlag. 1982.

Pertence J., A., "Amplificadores Operacionais e Filtros Ativos: Teoria, Projeto, Aplicações e Laboratório". 3o Edição. Editora Mac Graw Hill, São Paulo. 1988.

Powell, M. J. D., "Restart procedures for the conjugated gradient method", Mathematical Programming 12, pp. 241-254. April 1977.

Press, W. H.; Flannery, B. P.; Vetterling, W. T.; Numerical Recipes. The Art of Scientific Computing. Cambridge University Press, New York, 1990.

Qi, G.Z., Chen, H. M., Tsai, K. H., Yang, J. C. S., "Structural Dynamic Model Identification through Neural Network", ISMA19 - Tools for Noise and Vibration Analysis, 1994.

Roark, R. J.; "Formulas for Stress and Strain", McGraw-Hill, New York, 1965.

Roberts, J. B.; "Vibration Control of Flexible Transmission Shaft by Means of a Squeeze-Film Damper: An Experimental Investigation", Journal of Sound and Vibration, 113 (2), p. 329 - 345, 1987.

Romberg, T. M., Black, J. L., Ledwidge, T. J. "Signal Processing for Industrial Diagnostics", Series in Measurements Science and Technology. Editora John Wiley & Sons Ltd., Chichester, England, 1996.

Ruhl, R. e Booker, J.; "A Finite Element Model for Distributed Parameter Turbo-rotor Systems, ASME Journal of Eng. for Industry. 126-132, Feb. 1972.

Rumelhart, D. E.; Hinton, G. E. e Williams, R. J., "Learning representations by back-propagations errors", *Nature* 323. pp. 533 - 536. 1986.

Ruzzene , M. , Fasana, A., Garibaldi, I. and Piombo, B.."Natural Frequencies and Dampings Identification using Wavelet Transform: Aplication to Real Data",*Mechanical System and Signal Processing*, 1997, volume 11(2), pag 207-218.

Schimit, L. A.; *Structural Disign by Systematic Synthesis*, 2ª Conferência de Procedimentos em Computação Eletrônica, ASCE, New York, pp. 105-122 ,1960.

Seibold, S., Ben-Haim, Y., "Some Results on the Robust Reliability of Multi_hypotehesis Kalman Filters for Crack Diagnostics in Rotors", ISMA21,International Conference in Noise and Vibration Engineering. Leuven, Belgium. 1996.

Silva, F. M. e Almeida, L. B., "Speeding up back-propagation", *Advanced Neural Computers*, R. Eckmiller, ed., North-Holland. pp. 151-160. 1990.

Steffen Jr., V.; "Estudo Analítico e Experimental de Dinâmica de Rotores". *Revista Brasileira de Ciências. Mecânicas.* 3(3): 3-8 Set. 1981.

Steffen Jr., V.; Faria, M. L. M.; "Otimização de projetos Mecânicos", *Simpósio Mineiro de Mecânica Computacional*, UFMG, Belo Horizonte, pp. 415 - 428. Dezembro de 1991.

Uhl, Tadeusz, "Automation of Monitoring and Diagnostic Process for Rotating Machinery Using Neural Networks", ISMA19 - Tools for Noise and Vibration Analysis.

Vanderplaats, G.; *Numerical Optimization Techniques for Engineering Design*, McGraw-Hill, 1984.

Welstead, Stephen T., "Neural Network and Fuzzy: logic and applications in C/C++"; Wiley and Sons. New York 1996.

Werbos, P. J.; "Back-propagation Through Time: What it does and How to do it"; *IEEE*, vol. 78, no 10, pp. 1550 - 1560. Oct. 1990.

Windrow, B. e Sterns, S. D., "Adaptative Signal Processing", New York. Prentice - Hall, 1985.

Zurada, Jacek; Marks, Robert J.; Robinson, Charles J.; "Computational Intelligence Imitating Life"; IEEE Press. June 1994.

ANEXO I

Componentes da Bancada

INTENCIONALMENTE EM BRANCO

ANEXO II

Modelo de Elementos Finitos do Rotor - ANSYS

INTENCIONALMENTE EM BRANCO

ANEXO III

Modelo de Elementos Finitos da Estrutura da Bancada - ANSYS

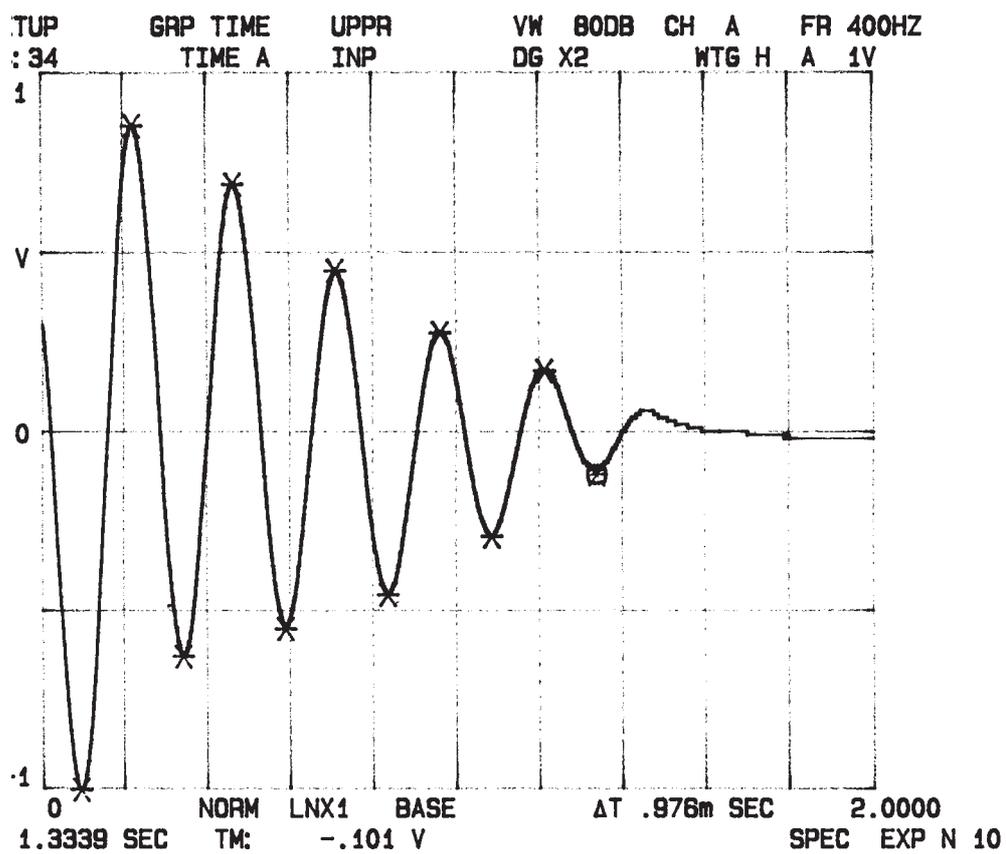
INTENCIONALMENTE EM BRANCO

ANEXO IV

Sistema de 1 GDL

Caracterização do Amortecedor

Resposta da montagem de 1 GDL à condição inicial de deslocamento.

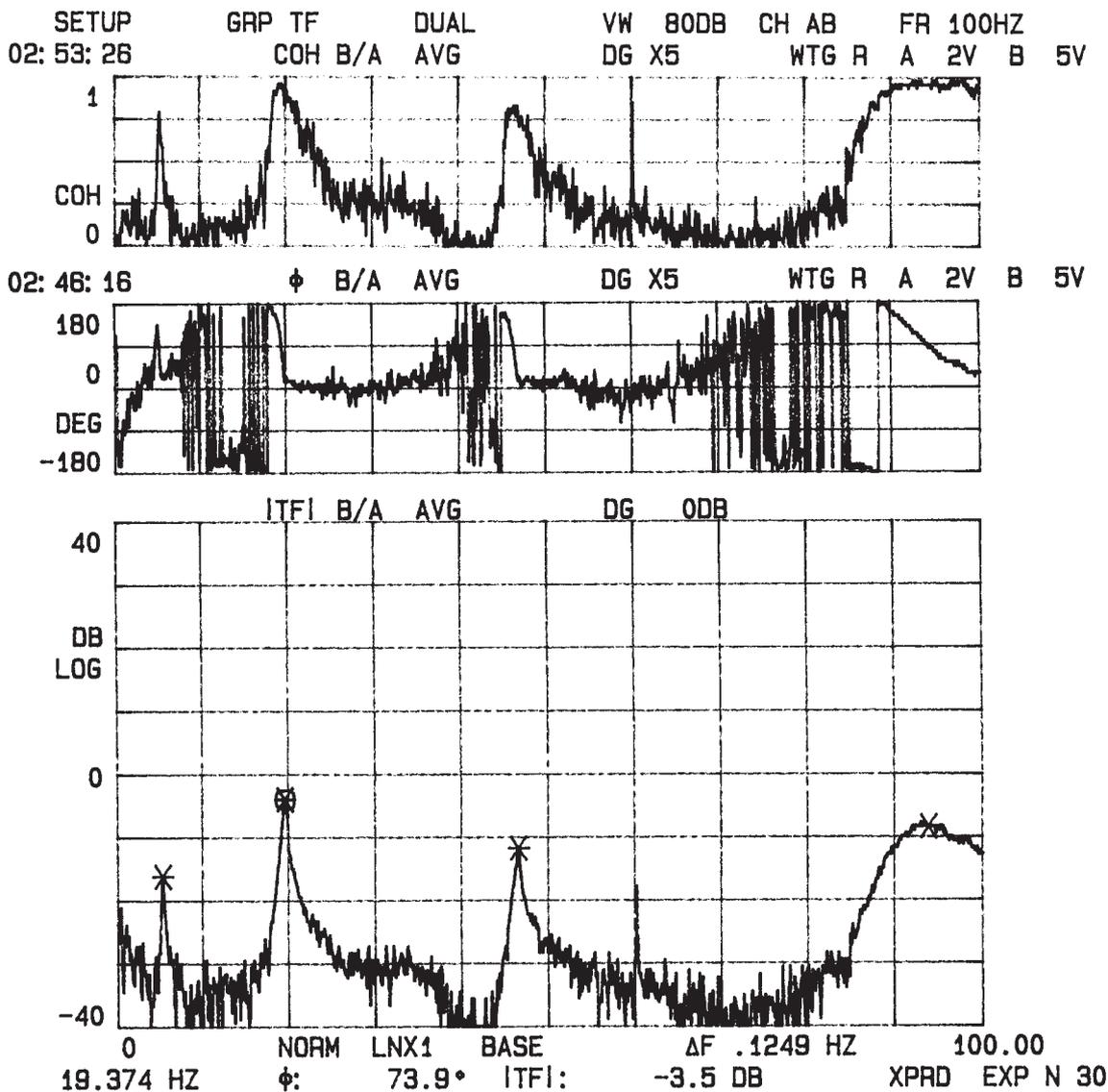


MARK	LIST X	Y
0	.091796	-.997
1	.21874	.861
2	.33788	-.617
3	.48093	.698
4	.58495	-.534
5	.70703	.453
6	.83105	-.442
7	.95702	.289

ANEXO V

Função de transferência do rotor

Rotação nula. Excitação ponto 2, medição ponto 4, como visto na figura (6.1).

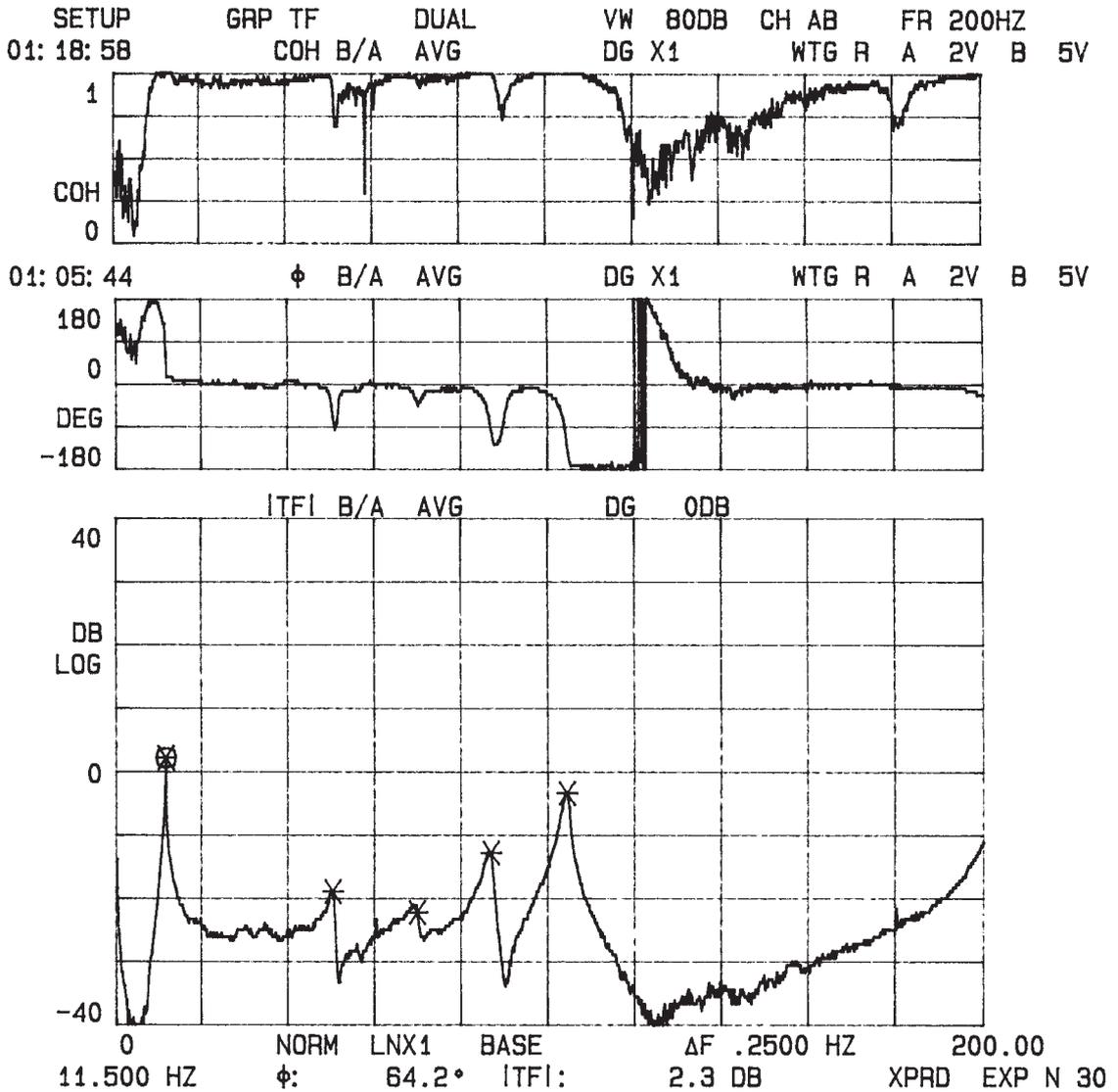


MARK	LIST X	φ	B/A	ITFI	B/A
0	5.2500	49.2		-15.5	
1	19.374	73.9		-3.5	
2	46.374	49.0		-11.7	
3	93.750	91.3		-7.7	

ANEXO VI

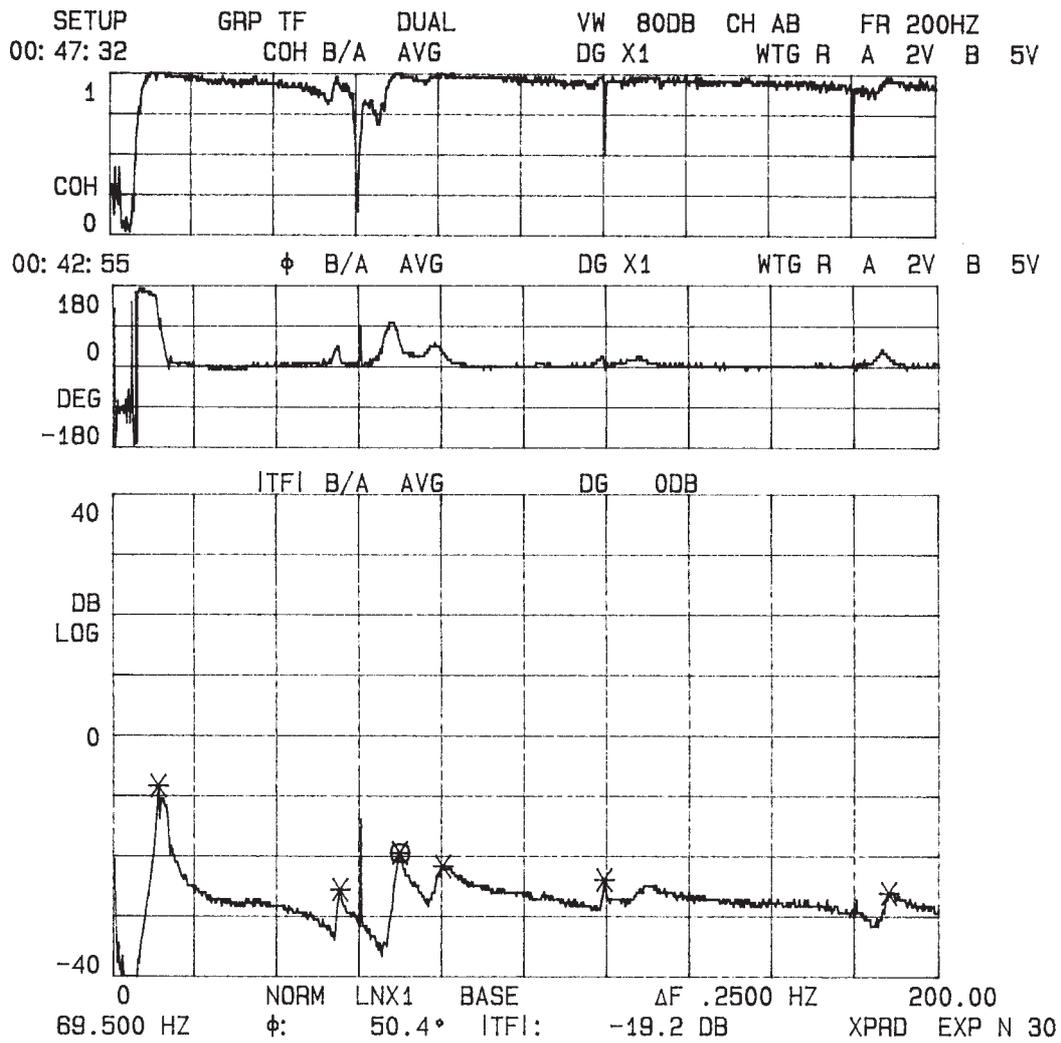
Função de transferência da estrutura da bancada.

Função de transferência e coerência, entre os pontos A₁A₂, como mostra a figura (5.7).

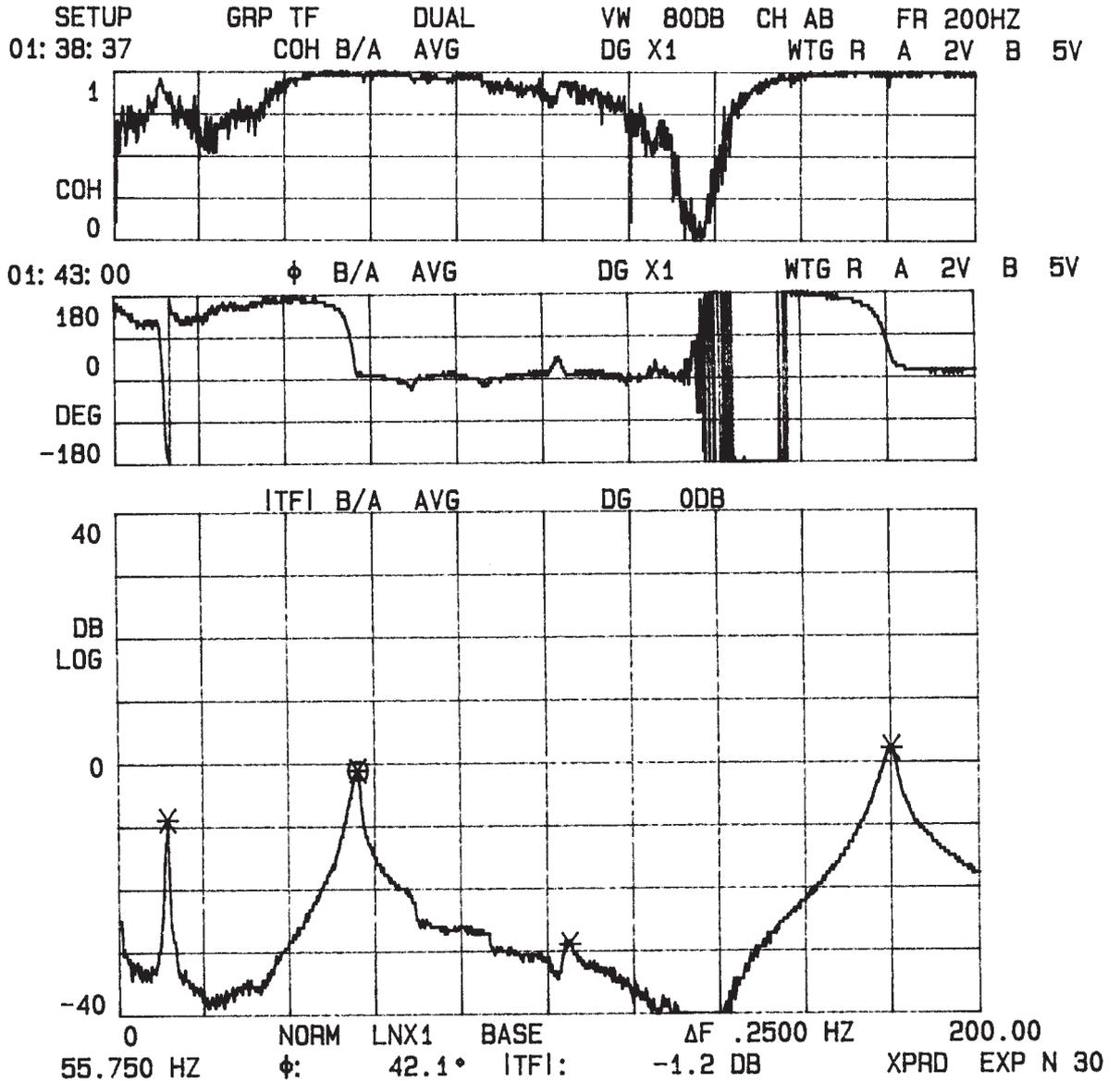


MARK	LIST X	ϕ B/A	ITFI B/A
0	11.500	64.9	2.5
1	50.000	-61.7	-18.7
2	69.500	-39.1	-21.8
3	86.500	-92.2	-12.8
4	104.00	-122.0	-2.8

Função de transferência e coerência, entre os pontos B₁B₂, como mostra a figura (5.7).



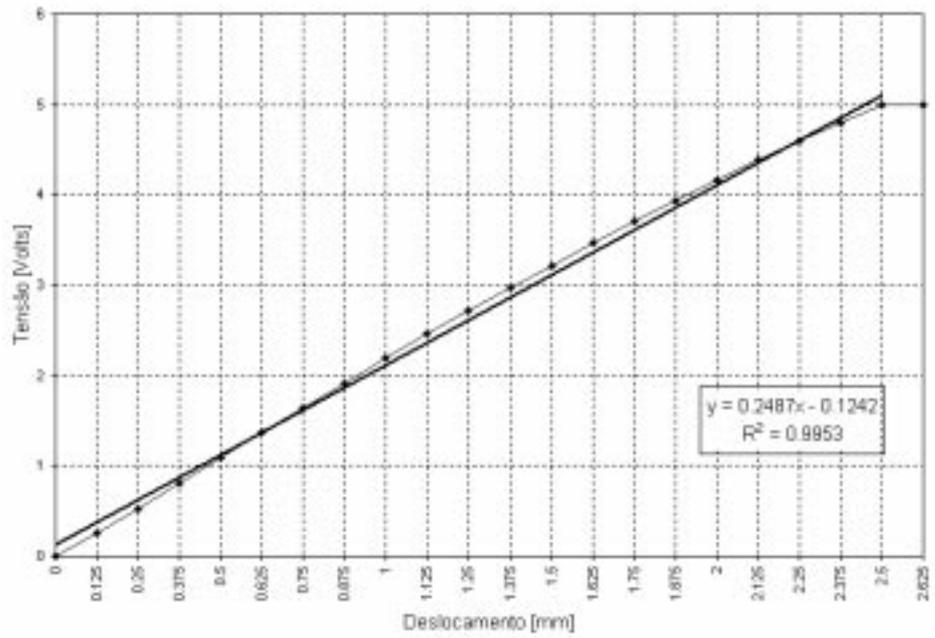
Função de transferência e coerência, entre os pontos C₁C₂, como mostra a figura (5.7).



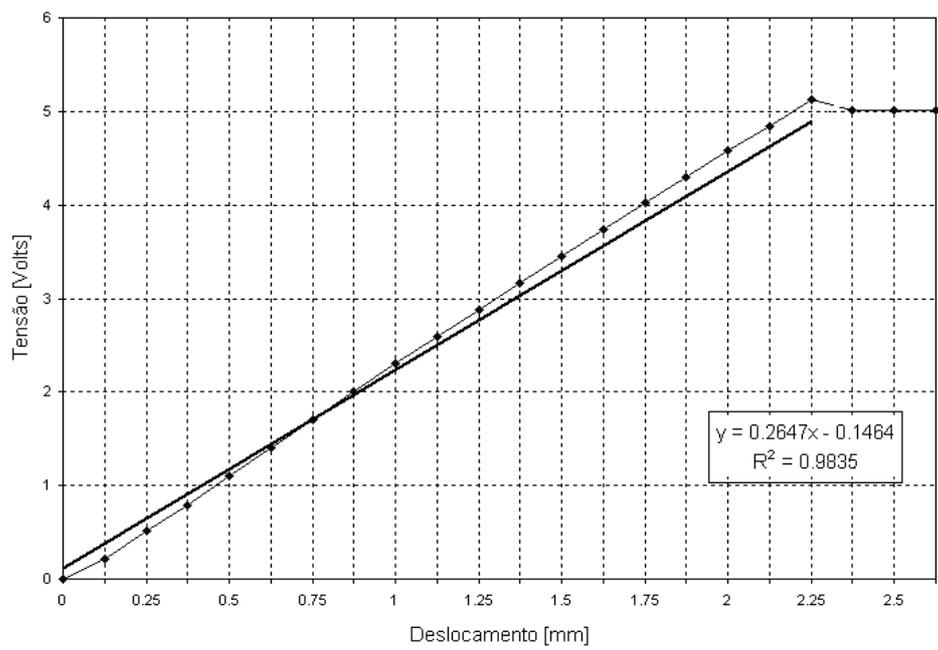
ANEXO VII

**Curvas de calibração do sistema de medição:
sensores de deslocamento (DYMAC)**

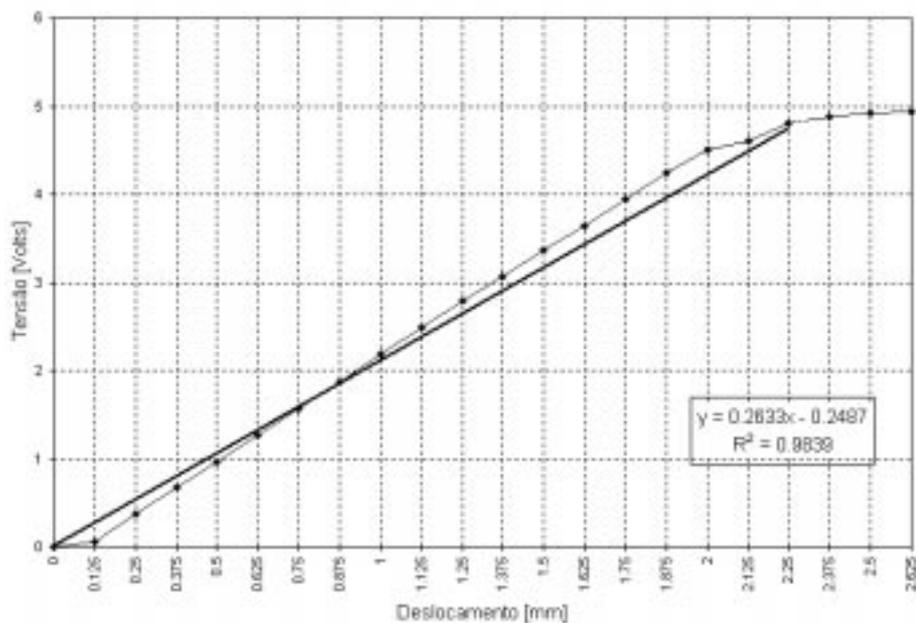
Calibração DYMAC 1



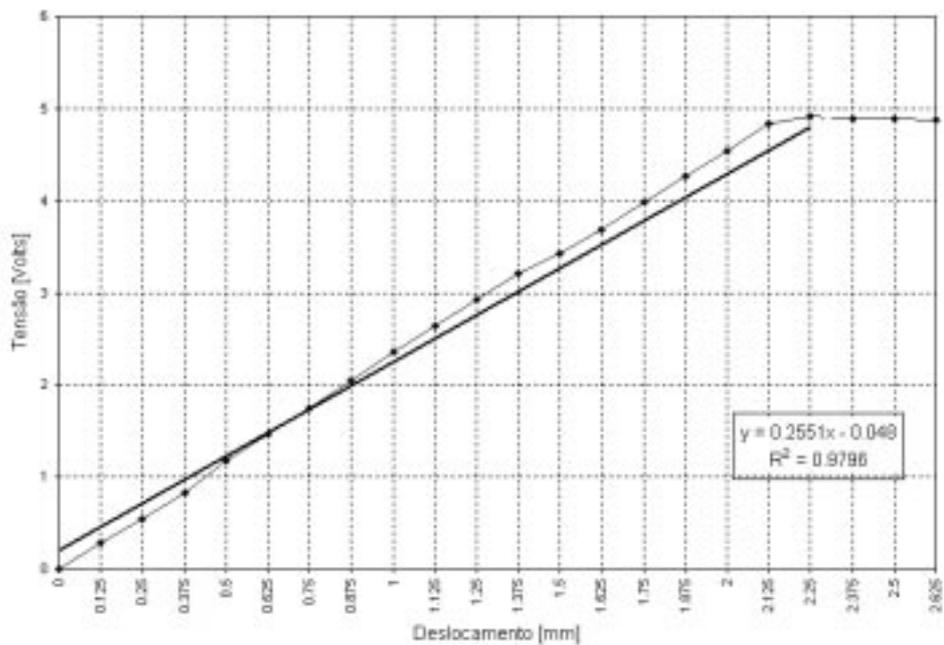
Calibração DYMAC 2

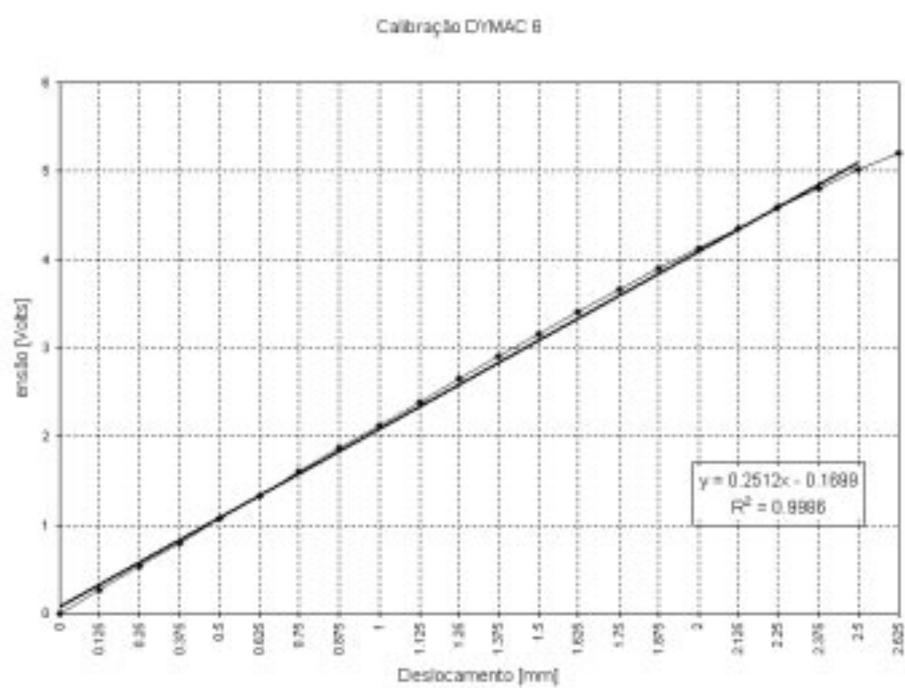
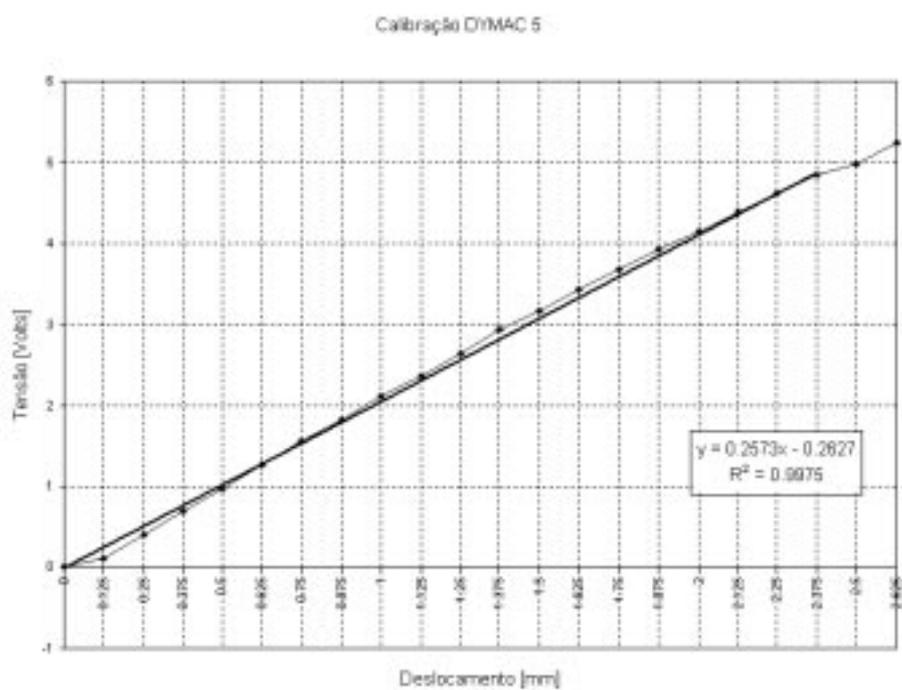


Calibração DYMAC 3



Calibração DYMAC 4





ANEXO VIII

Programa MKDX.for
Programa CUTGDL.for

INTENCIONALMENTE EM BRANCO

ANEXO IX

Programa de Rede Neural

(BACKPMA.C)

INTENCIONALMENTE EM BRANCO

ANEXO X

Programa COMPACT.M

INTENCIONALMENTE EM BRANCO