

UNIVERSIDADE FEDERAL DE SANTA CATARINA

CURSO DE PÓS-GRADUAÇÃO EM ENGENHARIA DE PRODUÇÃO E SISTEMAS

MARCO ANDRÉ MAZZAROTTO

ANÁLISE NÃO-LINEAR DE SÉRIES TEMPORAIS UNIVARIADAS
Modelagem, previsão e caos

Tese de doutorado apresentada como requisito parcial à obtenção do grau de Doutor em Engenharia de Produção. Curso de Engenharia de Produção e Sistemas. Universidade Federal de Santa Catarina. Orientador: Prof. Robert Wayne Samohyl, Ph.D.

FLORIANÓPOLIS

1996

ANÁLISE NÃO-LINEAR DE SÉRIES TEMPORAIS UNIVARIADAS
Modelagem, Previsão e Caos

MARCO ANDRÉ MAZZAROTTO

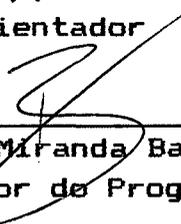
Esta tese foi julgada adequada para a obtenção do título de

DOUTOR EM ENGENHARIA

especialidade Engenharia de Produção e aprovada em sua forma final pelo programa de Pós-Graduação em Engenharia de Produção e Sistemas da Universidade Federal de Santa Catarina.

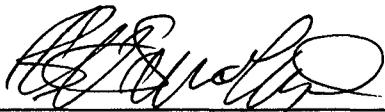


Prof. Robert Wayne Samohyl, Ph.D.
Orientador

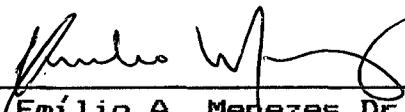


Prof. Ricardo Miranda Barcia, Ph.D.
Coordenador do Programa

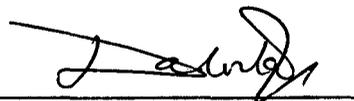
BANCA EXAMINADORA



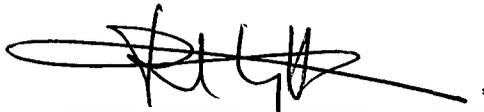
Prof. Robert W. Samohyl, Ph.D.
Orientador



Prof. Emílio A. Menezes, Dr.



Prof. Newton C.A. da Costa Jr., Dr.
Examinador Externo



Prof. Ricardo L. Viana, Dr.
Examinador Externo



Prof.ª Eliza Z. Toma, Dr.ª.
Moderadora

JUSTIFICATIVA

O professor Dr. Plínio Stange, membro da banca examinadora para apresentação e defesa desta tese de doutorado faleceu, prematuramente, em 13/10/96.

AGRADECIMENTOS

Ao professor Robert Wayne Samohyl, pela orientação desta tese.

Ao professor Newton A.C. da Costa Jr., pela participação decisiva na escolha do tema desta tese.

Ao professor Plínio Stange pelo apoio em todas as etapas da realização deste curso de doutorado.

À Universidade Federal do Paraná e, em particular, ao Departamento de Matemática pela licença concedida para a realização do curso de doutorado.

À CAPES, Coordenação de Aperfeiçoamento de Pessoal de Nível Superior, pelo apoio financeiro.

SUMÁRIO

<u>TERMO DE APROVAÇÃO</u>	i
<u>JUSTIFICATIVA</u>	ii
<u>AGRADECIMENTOS</u>	iii
<u>LISTA DE QUADROS</u>	ix
<u>LISTA DE FIGURAS</u>	xi
<u>LISTA DE ABREVIATURAS</u>	xiv
<u>RESUMO</u>	xv
<u>ABSTRACT</u>	xvi
<u>INTRODUÇÃO</u>	1

CAPÍTULO 1

SISTEMAS DETERMINÍSTICOS CAÓTICOS

1.1 CARACTERÍSTICAS DOS SISTEMAS DETERMINÍSTICOS CAÓTICOS.....	5
1.2 SISTEMAS DETERMINÍSTICOS CAÓTICOS CLÁSSICOS.....	6
1.2.1 Tent.....	6
1.2.2 Logística.....	8
1.2.3 Hénon.....	9
1.2.4 A equação de Mackey-Glass discreta.....	9
1.2.5 Ikeda.....	10
1.2.6 O modelo de Lorenz.....	10
1.2.7 Caos de alta dimensão.....	11
1.2.8 Séries não-lineares com ruído.....	12
1.3 GEOMETRIA DO CAOS.....	12
1.4 EXPOENTE DE LYAPUNOV.....	13

1.5 PREVISÃO.....	16
1.6 ATRADORES.....	16
1.6.1 Conceito.....	16
1.6.2 Reconstrução do atrator e da dinâmica de um sistema.....	17
O teorema de Takens.....	17
Reconstrução.....	18
1.6.3 Exemplos de reconstrução.....	20
O atrator de Hénon.....	20
O atrator de Mackey-Glass.....	21

CAPÍTULO 2

IDENTIFICAÇÃO DE SISTEMAS DETERMINÍSTICOS CAÓTICOS

Métodos clássicos

2.1 DIMENSÃO.....	23
2.1.1 Capacidade.....	23
2.1.2 Dimensão de correlação.....	27
2.1.3 Algoritmos.....	28
2.1.4 Distinguindo sistemas determinísticos caóticos de ruído.....	30
2.1.5 Exemplos numéricos.....	31
2.1.6 Limitações.....	34
2.2 EXPOENTES DE LYAPUNOV.....	40
2.2.1 Definições.....	40
2.2.2 Sistemas dinâmicos não-lineares.....	42
2.2.3 Algoritmos para estimar o expoente de Lyapunov dominante.....	43
2.2.4 Exemplos numéricos.....	44
2.2.5 Limitações.....	47

CAPÍTULO 3

PREVISÃO DE SÉRIES TEMPORAIS NÃO-LINEARES

3.1 MODELOS DE PREVISÃO.....	49
3.1.1 Tipos de modelos.....	50
Modelos locais.....	50
Modelos globais.....	50
3.2 REGRESSÃO NÃO-PARAMÉTRICA LOCAL.....	51
3.2.1 Nearest-neighbor.....	51

3.2.2	Locally-weighted regression.....	53
3.3	ESTATÍSTICAS.....	54
3.3.1	Erro quadrático médio.....	54
3.3.2	Estatística BDS.....	55
3.4	EXEMPLOS NUMÉRICOS.....	56
3.5	REGRESSÃO NÃO-PARAMÉTRICA GLOBAL.....	61
3.5.1	Redes neurais.....	61
3.5.2	Função de ativação.....	63
3.5.3	Estimação da rede neural.....	64
3.5.4	Exemplos numéricos.....	65
	Séries determinísticas caóticas e séries não-lineares com ruído.....	65
	A série das Manchas Solares.....	66

CAPÍTULO 4

MODELAGEM DE SÉRIES TEMPORAIS NÃO-LINEARES

4.1	MODELAGEM.....	72
4.1.1	O algoritmo de modelagem.....	73
4.2	AS CURVAS EQMN(T,m,p,k).....	74
4.2.1	Modelagem determinística vs. modelagem estocástica.....	74
4.2.2	Dimensão do modelo.....	75
4.2.3	Previsão de curto prazo.....	75
4.3	RESULTADOS NUMÉRICOS.....	76
4.3.1	Modelagem determinística vs. modelagem estocástica.....	77
4.3.2	Dimensão do modelo.....	79
4.3.3	Previsão de curto prazo.....	81
4.4	DIAGNOSTICANDO CAOS EM SÉRIES TEMPORAIS.....	83
4.5	TESTES ESTATÍSTICOS.....	89
4.5.1	Sistemas não-lineares com ruído.....	89
4.5.2	Sistemas determinísticos caóticos.....	90

CAPÍTULO 5

NÃO-LINEARIDADES EM SÉRIES FINANCEIRAS BRASILEIRAS

5.1	MODELAGEM.....	94
5.1.1	Modelagem linear vs. modelagem não-linear.....	95
5.1.2	Dimensão do modelo.....	95

5.1.3	Testes estatísticos.....	96
	Teste do terceiro momento incondicional.....	98
	Teste de McLeod-Li.....	98
	Teste BDS.....	98
5.1.4	Dimensão de correlação.....	99
5.2	O MODELO ARCH(p).....	101
5.2.1	Definição.....	102
5.2.2	Identificação e checagem.....	102
5.2.3	Estimação dos modelos ARCH(p).....	103
5.3	MODELOS ARCH(p) ESTIMADOS.....	103
5.3.1	Coefficientes e erros padrões.....	103
5.3.2	Testes estatísticos.....	104
	Teste de McLeod-Li.....	104
	Teste BDS.....	106
5.3.3	Dimensão de correlação.....	106
 <u>CONCLUSÃO</u>		 109
 <u>RECOMENDAÇÕES</u>		 112

APÊNDICE A

TESTES ESTATÍSTICOS

A.1	TESTE ESTATÍSTICO BDS.....	113
A.2	TESTE DO TERCEIRO MOMENTO INCONDICIONAL.....	115
	A.2.1 Modelos não-lineares.....	115
	A.2.2 O teste do terceiro momento incondicional.....	116

APÊNDICE B

TEOREMA DE TAKENS

B.1	SISTEMA DINÂMICO DE DIMENSÃO 2.....	118
B.2	CIRCUNFERÊNCIA UNITÁRIA.....	119
B.3	HÉNON.....	120

APÊNDICE C

ENTROPIA MÉTRICA E EXPOENTE DE LYAPUNOV DOMINANTE

C.1 REGRESSÃO NÃO-PARAMÉTRICA LWR.....122
C.2 REGRESSÃO NÃO-PARAMÉTRICA NN.....125

REFERÊNCIAS BIBLIOGRÁFICAS.....127

LISTA DE QUADROS

2.1-DIMENSÃO DE CORRELAÇÃO PARA OS PROCESSOS ESTOCÁSTICOS NORMAL IID E ARCH(2).....	32
2.2-EXPOENTE DE LYAPUNOV DOMINANTE. SISTEMAS DETERMINÍSTICOS CAÓTICOS.....	45
2.3-EXPOENTE DE LYAPUNOV DOMINANTE. SISTEMAS NÃO-LINEARES COM RUÍDO.....	46
3.1-ERRO QUADRÁTICO MÉDIO NORMALIZADO-EQMN(1). REGRESSÃO NÃO-PARAMÉTRICA NN E LWR.....	56
3.2-ERRO QUADRÁTICO MÉDIO NORMALIZADO-EQMN(1). REGRESSÃO NÃO-PARAMÉTRICA LWR COM ESQUEMA DE PESOS UNIFORME E TRICÚBICA.....	57
3.3-REDES NEURAIS. PARÂMETROS N E m , ERRO QUADRÁTICO MÉDIO EQM.....	69
3.4.a-COEFICIENTES E ERROS PADRÕES DA REDE NEURAL AJUSTADA À SÉRIE DAS MANCHAS SOLARES(Continua no quadro 3.4.b).....	69
3.4.b-COEFICIENTES E ERROS PADRÕES DA REDE NEURAL AJUSTADA À SÉRIE DAS MANCHAS SOLARES(Continuação do quadro 3.4.a)...	70
3.5-ESTATÍSTICA BDS PARA OS RESÍDUOS DA REDE NEURAL AJUSTADA À SÉRIE DAS MANCHAS SOLARES.....	70
3.6-TESTE DO TERCEIRO MOMENTO INCONDICIONAL PARA A SÉRIE DAS MANCHAS SOLARES.....	71
4.1-EXPOENTES DE LYAPUNOV DOMINANTE PARA OS SISTEMAS AUTO- REGRESSIVOS.....	79

4.2-VALORES ÓTIMOS DOS PARÂMETROS m E k QUE MINIMIZAM O ERRO QUADRÁTICO MÉDIO $EQMN(m,k)$	81
4.3.a- $EQMN(p)$ PARA OS SISTEMAS DO GRUPO I: SISTEMAS DETERMINÍSTICOS CAÓTICOS.....	88
4.3.b- $EQMN(p)$ PARA OS SISTEMAS DOS GRUPOS II E III: SISTEMAS ESTOCÁSTICOS E SISTEMAS NÃO-LINEARES COM RUÍDO.....	88
4.4-TESTE DO TERCEIRO MOMENTO PARA OS MODELOS LOCAIS DE HÉNON6 E HÉNON22.....	91
4.5-ESTATÍSTICA BDS(1.0 DP) PARA OS MODELOS LOCAIS DE HÉNON6 E HÉNON22.....	91
4.6-TESTE DO TERCEIRO MOMENTO INCONDICIONAL PARA OS MODELOS LOCAIS DE LORENZ1.5 E LORENZ12.....	92
4.7- ESTATÍSTICA BDS(1.0 DP) PARA OS MODELOS LOCAIS DE LORENZ1.5 E LORENZ12.....	92
4.8-TESTE DO TERCEIRO MOMENTO INCONDICIONAL PARA OS MODELOS LOCAIS DE HÉNON E LORENZ.....	93
4.9-ESTATÍSTICA BDS(1.0 DP) PARA OS RESÍDUOS DOS MODELOS LOCAIS DE HÉNON E LORENZ FILTRADOS POR MODELO LINEAR.....	93
5.1-VALORES ÓTIMOS DOS PARÂMETROS m E k QUE MINIMIZAM $EQMN(m,k)$	96
5.2-ESTATÍSTICAS BÁSICAS DAS SÉRIES FINANCEIRAS.....	96
5.3-TESTE DO TERCEIRO MOMENTO INCONDICIONAL PARA OS RESÍDUOS DOS MODELOS LINEARES.....	100
5.4-TESTE DE PORTEMATAU(TP) E DE MCLEOD-LI(TML) PARA AS SÉRIES ORIGINAIS E PARA OS RESÍDUOS DOS MODELOS LINEARES..	100
5.5-ESTATÍSTICA BDS(1.0 DP) PARA OS RESÍDUOS DOS MODELOS LINEARES.....	101
5.6-DECLIVIDADES DAS RETAS AJUSTADAS ÀS DIMENSÕES DE CORRELAÇÃO DOS RESÍDUOS DOS MODELOS LINEARES.....	101

5.7-COEFICIENTES E ERROS PADRÕES DOS MODELOS ARCH(p) ESTIMADOS PARA AS SÉRIES FINANCEIRAS.....	104
5.8-ESTATÍSTICAS BÁSICAS DOS RESÍDUOS PADRONIZADOS DOS MODELOS ARCH(p) ESTIMADOS PARA AS SÉRIES FINANCEIRAS.....	104
5.9-TESTE DE MCLEOD AND LI(TML) PARA OS RESÍDUOS PADRONIZADOS DOS ARCH(p).....	105
5.10-ESTATÍSTICA BDS(1.0 DP) PARA OS RESÍDUOS PADRONIZADOS DOS ARCH(p).....	106
C.1-ENTROPIA MÉTRICA h. REGRESSÃO NÃO-PARAMÉTRICA LWR.....	124
C.2-EXPOENTE DE LYAPUNOV DOMINANTE. REGRESSÃO NÃO- PARAMÉTRICA NN	125

LISTA DE FIGURAS

1.1.a Órbita(trajetória) do sistema determinístico tent.....	7
1.1.b Separação das trajetórias causada pela sensibilidade às condições iniciais do sistema dinâmico tent.....	7
1.2.a Mecanismos geométricos do sistema dinâmico tent.....	12
1.2.b Representação da ação dos mecanismos geométricos dos sistemas determinísticos caóticos sobre um elemento de volume arbitrário.....	13
1.3 Separação das condições iniciais x_0 e x_0^* , pelo sistema dinâmico $x_{t+1}=f(x_t)$	14
1.4 Trajetórias da Logística com $\alpha=3.568$, obtidas das condições iniciais $x_0=0.3$ e $x_0^*=0.305$	15
1.5.a Atrator caótico do sistema dinâmico de Hénon.....	21
1.5.b Atrator caótico do sistema dinâmico de Mackey-Glass.....	22

2.1.a	Recobrimento de um ponto isolado por cubos de dimensão 2 e aresta ϵ	24
2.1.b	Recobrimento de uma curva C de comprimento 1 por cubos de dimensão 2 e aresta ϵ	25
2.1.c	Recobrimento do atrator do sistema dinâmico de Hénon por cubos de dimensão 2 e aresta ϵ	26
2.2	Gráficos das curvas $\log C(T, \epsilon, m)$ vs. $\log \epsilon$, para o sistema dinâmico de Mackey-Glass (fig. 2.2.a). Gráfico da dimensão de correlação $\nu(T, m)$, para o modelo de Lorenz (fig. 2.2.b).....	35
2.3	Gráficos das curvas $\log C(T, \epsilon, m)$ vs. $\log \epsilon$, para o modelo de Lorenz (fig. 2.3.a), Lorenz1.5 (fig. 2.3.b) e Lorenz12 (fig. 2.3.c).....	36
2.4	Gráficos das curvas $\log C(T, \epsilon, m)$ vs. $\log \epsilon$, para o sistema dinâmico de Hénon (fig. 2.4.a), Hénon6 (fig. 2.4.b) e Hénon22 (fig. 2.4.c).....	37
2.5	Gráficos da dimensão de correlação $\nu(T, \epsilon, m)$, para o modelo de Lorenz (fig. 2.5.a), Lorenz1.5 (fig. 2.5.b) e Lorenz12 (fig. 2.5.c).....	38
2.6	Gráficos das curvas $\log C(T, \epsilon, m)$ vs. $\log \epsilon$, para o processo normal iid. (fig. 2.6.a) e ARCH(2) (fig. 2.6.b).....	39
3.1	Comparação dos valores observados x_{t+1} com os valores das previsões \hat{x}_{t+1} , $t=501, \dots, 549$, para o sistema dinâmico de Hénon (fig. 3.1.a), Hénon6 (fig. 3.1.b) e Hénon22 (fig. 3.1.c).....	58
3.2	Comparação dos valores observados x_{t+1} com os valores das previsões \hat{x}_{t+1} , $t=501, \dots, 549$, para o modelo de Lorenz (fig. 3.2.a), Lorenz1.5 (fig. 3.2.b) e Lorenz12 (fig. 3.2.c).....	59
3.3	Comparação dos valores observados x_{t+1} com os valores das previsões \hat{x}_{t+1} , $t=231, \dots, 279$, para a série das Manchas Solares. Regressão não-paramétrica LWR e NN.....	60

3.4	Representação esquemática do elemento básico da rede neural (3.20), o neurônio.....	62
3.5	Arquitetura da rede neural (3.20).....	62
3.6	Exemplos de funções de ativação(transferência) para a rede neural(3.20).....	64
3.7	Comparação dos valores observados x_{t+1} com os valores \hat{x}_{t+1} dados pela rede neural (3.20), ajustada à série das Manchas Solares, $t=4, \dots, 279$	68
3.8	Comparação dos valores observados x_{t+1} com os valores \hat{x}_{t+1} , $t=229, \dots, 279$, dados pela rede neural ajustada à série das Manchas Solares.....	60
3.9	Dimensão de correlação para a série das Manchas Solares e dos resíduos da rede neural (3.20) ajustada à série; dimensão de correlação da lei de escala $\nu(m)=m$	68
4.1	Gráficos das curvas EQMN(k), $k=20, \dots, 420$, para os sistemas dinâmicos do grupo I(Hénon, Ikeda, Lorenz e Mackey-Glass) (fig. 4.1.a), grupo II(AR(2), ARCH(2) e Manchas Solares) (fig. 4.1.b) e grupo III(Hénon6, Hénon22, Lorenz1.5 e Lorenz12(fig. 4.1.c).....	84
4.2	Ampliação da figura 4.2, para os modelos AR(2) e Manchas Solares(fig. 4.2.a), e para os sistemas do grupo III(Hénon6, Hénon22, Lorenz1.5 e Lorenz12)(fig. 4.2.b).....	85
4.3	Gráficos das curvas EQMN(m), $m=1, \dots, 10$, para os sistemas dinâmicos do grupo I(Hénon, Ikeda, Lorenz e Mackey-Glass) (fig. 4.3.a), grupo II(AR(2), ARCH(2) e Manchas Solares) (fig. 4.3.b) e grupo III(Hénon6, Hénon22, Lorenz1.5 e Lorenz12(fig. 4.3.c).....	86
4.4	Gráficos das curvas EQMN(p), $p=1, \dots, 10$, para os sistemas dinâmicos do grupo I(Logística, Hénon, Ikeda, Mackey-Glass), obtidas através da regressão não-paramétrica LWR.....	87
4.5	Gráficos das curvas EQMN(T), $T_1=400, 600, \dots, 1200$, $T_2=300$, para os sistemas dinâmicos de Hénon e Lorenz.....	87

5.1 Gráficos das curvas EQMN(k), k=20,...,460, para as séries financeiras(retornos) BRASFUND, WHM e VALE.....	97
5.2 Gráficos das curvas EQMN(m), m=1,...,10, para as séries financeiras(retornos) BRASFUND, WHM e VALE.....	97
5.3 Dimensão de correlação para os resíduos dos modelos lineares ajustados às séries financeiras(retornos) IBOVESPA, BRASFUND, WHM e VALE; dimensão de correlação da lei de escala $\nu(m)=m$	106
5.4 Dimensão de correlação para os resíduos dos modelos AR(5) e ARCH(5) ajustados à série financeira(retornos) IBOVESPA; dimensão de correlação da lei de escala $\nu(m)=m$..	106

LISTA DE ABREVIATURAS

As abreviaturas

1. BRASFUND
2. VALE
3. WHM
4. IBOVESPA

referem-se, respectivamente, aos retornos das séries financeiras brasileiras: títulos brasileiros negociados na Bolsa de Nova Iorque(1), preços de fechamento das ações das empresas Vale do Rio Doce(2) e White Martins(3), e do índice da Bolsa de Valores de São Paulo(4).

5. $N(0, \sigma^2)$ - Processo normal com média 0.0 e variância σ^2
6. iid. - Processo independente e identicamente distribuído
7. AR(p) - Modelo estocástico auto-regressivo linear de ordem p
8. ARCH(p) - Modelo estocástico auto-regressivo, com variância condicional(heterocedástica), de ordem p

RESUMO

Este trabalho trata da análise de não-linearidades em séries temporais, utilizando técnicas da dinâmica não-linear e das regressões não-paramétricas. Testes estatísticos conhecidos são empregados para identificar não-linearidades e se as não-linearidades observadas podem ser explicadas por modelos estocásticos não-lineares ou por modelos determinísticos caóticos. Modelos de previsão de séries não-lineares desenvolvidos na Física, implementados neste trabalho através de regressões não-paramétricas disponíveis na literatura econométrica, são utilizados para decidir qual das três modelagens, a linear estocástica, a não-linear auto-regressiva estocástica ou a não-linear determinística, é a mais apropriada para certas séries temporais. As idéias, técnicas e métodos descritos são aplicados a séries simuladas artificialmente e a uma série do mundo real. Finalmente, não-linearidades são analisadas em séries financeiras brasileiras, procurando-se a origem das não-linearidades observadas e suas relações com as técnicas e modelos aqui apresentados.

ABSTRACT

The present work intends to analyse the non-linearities found in time series making use of non-linear dynamic techniques and of nonparametric regressions. Some well known statistics tests are managed to identify non-linearities and verify if these non-linearities could be explained by the non-linear stochastic models or by the chaotic deterministic models. Forecast models of non-linear series developed by the Physics and implemented in the present work through the nonparametric regressions available in the econometric literature are used to define wich modelling - the stochastic linear, the stochastic non-linear autoregressive or the deterministic non-linear - is the most appropriated one to determined time series. The ideas, techniques and methods here described are applicable as to artificially simulated series as to a real world serie. Finally, the non-linearities are analysed in certain brazilian financial time series as to search the origin of those non-linearities and the way they are related to the techniques and models employed.

INTRODUÇÃO

Nos últimos anos, muitos trabalhos em séries temporais têm investigado caos determinístico em séries do mundo real. Em particular, em Economia e Finanças foram depositadas esperanças de que os modelos determinísticos caóticos fossem capazes de explicar as grandes flutuações e o complexo equilíbrio de longo prazo dos sistemas dinâmicos geradores de algumas dessas séries.

Dois pontos em comum podem ser destacados nesses trabalhos:

- 1) a investigação de caos determinístico através da dimensão de correlação;
- 2) a idéia de que a presença de caos justificaria a imprevisibilidade do comportamento futuro da série temporal a partir de dados passados observados.

A dimensão de correlação, técnica com origens na Física, empregada para investigar caos em séries temporais com comprimento (número de observações) ilimitado e baixo nível de ruído, revelou-se inadequada às séries do mundo real, onde os sistemas dinâmicos geradores das séries não estão sob controle do pesquisador e têm em geral pequeno comprimento e nível de ruído significativo. Embora a dimensão de correlação tenha apontado fortes indícios de não-linearidades e a provável presença de caos em algumas séries, os resultados não são conclusivos devido ao emprego de algoritmos dessa técnica fora do contexto para o qual foram desenhados.

A idéia de que a presença de caos determinístico justificaria a imprevisibilidade da série temporal é criada no leitor desses trabalhos de forma subjetiva e não intencional, através da omissão de referências a técnicas de previsão para as séries, eventualmente identificadas como caóticas. Essa idéia, correta para previsões de longo prazo, não é verdadeira para previsões de curto prazo. Séries determinísticas caóticas são previsíveis a curto prazo. O erro e o horizonte da previsão dependem da sensibilidade às condições iniciais, do comprimento e do nível de ruído da série observada do sistema dinâmico.

O resultado teórico, devido a Takens, o qual permite o emprego da dimensão de correlação para investigar caos determinístico em sistemas dinâmicos a partir da série univariada

observada, afirma também que sistemas determinísticos caóticos podem ser modelados por modelos auto-regressivos não-lineares, indicando outra forma de investigar caos determinístico, mais próxima da metodologia das Séries Temporais: a modelagem da série temporal por modelo não-linear auto-regressivo (modelo de previsão) e a investigação de quando o caos pode explodir nesses modelos.

Modelos de previsão de séries determinísticas caóticas surgiram na Física com o trabalho de Farmer and Sidorowich (1987). Com esses modelos, os autores investigaram caos em sistemas determinísticos caóticos, avaliando indiretamente a qualidade dos modelos através dos erros de previsão e da reprodução de invariantes geométricos e dinâmicos conhecidos dos sistemas. Casdagli (1992) retoma os modelos de previsão de Farmer and Sidorowich e mostra que é possível decidir sobre a modelagem mais apropriada para dada série temporal, com o auxílio desses modelos.

Uma característica dos modelos de previsão chama a atenção: o comprimento da série temporal para construir esses modelos é substancialmente menor que o exigido para estimar a dimensão de correlação, embora possa ser exagerado para séries do mundo real.

Motivado por essas colocações, pretende-se:

1) avaliar o desempenho dos modelos de previsão de Farmer and Sidorowich e do algoritmo de modelagem de Casdagli, implementados neste trabalho com regressões não-paramétricas disponíveis na literatura, em séries temporais não-lineares com pequeno comprimento e com nível de ruído significativo, com a avaliação indireta dos modelos de previsão substituída por testes estatísticos adequados aos resíduos dos modelos;

2) investigar não-linearidades em séries financeiras brasileiras e, em particular, investigar se essas não-linearidades podem ser explicadas pela existência de variáveis auto-regressivas não-lineares nos sistemas dinâmicos geradores dessas séries.

Este trabalho, buscando atingir os objetivos estabelecidos em 1) e 2), está organizado em cinco capítulos. No capítulo 1, sistemas determinísticos caóticos clássicos, são apresentados. Conceitos e características básicas da dinâmica não-linear caótica e o teorema de Takens, que permite a reconstrução do sistema dinâmico a partir da série temporal univariada observada, são introduzidos. A geometria dos sistemas determinísticos caóticos, que permite visualizar por que a previsão de séries determinísticas caóticas não pode ser feita de forma eficiente

com os métodos tradicionais de previsão, é enfatizada. As séries determinísticas caóticas observadas nesses sistemas são utilizadas em algoritmos clássicos da análise de sistemas dinâmicos não-lineares e para gerar séries não-lineares com ruído, empregadas para testar os algoritmos de previsão, modelagem e análise da sensibilidade às condições iniciais de sistemas estocásticos auto-regressivos não-lineares.

No capítulo 2, faz-se uma revisão de duas técnicas clássicas da análise de sistemas dinâmicos não-lineares: a dimensão de correlação e o expoente de Lyapunov. Algoritmos e exemplos numéricos para as séries do capítulo 1, são apresentados. Um conceito mais amplo de sensibilidade às condições iniciais, que permite analisar essa propriedade em sistemas estocásticos não-lineares, é apresentado. Esse conceito será utilizado para complementar os resultados obtidos com o algoritmo de modelagem no capítulo 4. As limitações dessas técnicas também são abordadas. As limitações não são, propriamente, às técnicas da dimensão de correlação e da análise da sensibilidade às condições iniciais, e sim a utilização de algoritmos dessas técnicas fora do contexto para o qual foram desenhados: séries determinísticas com grande comprimento e com baixo nível de ruído.

No capítulo 3, as regressões não-paramétricas NN(nearest-neighbor), LWR(locally-weighted regression) e redes neurais utilizadas para construir os modelos de previsão são apresentadas. Dois tipos de modelos de previsão são discutidos: modelos locais e modelos globais. Os modelos locais são obtidos das regressões não-paramétricas NN e LWR, enquanto os modelos globais são dados por redes neurais. Modelos locais e globais são construídos para séries determinísticas caóticas e para séries não-lineares com ruído, observadas dos sistemas determinísticos do capítulo 1. Os erros de previsão e gráficos que comparam os valores observados com os valores das previsões são utilizados qualitativamente para avaliar o desempenho desses modelos.

No capítulo 4, os modelos de previsão são empregados para modelar séries temporais não-lineares. O algoritmo de modelagem que permite decidir qual entre as três modelagens, a estocástica auto-regressiva linear, a estocástica auto-regressiva não-linear e a não-linear determinística, é aplicado em séries determinísticas caóticas e séries não-lineares com ruído obtidas de sistemas do capítulo 1; em séries lineares e não-lineares simuladas numericamente e a uma série não-linear do mundo real. Nos sistemas modelados de forma determinística ou por modelo auto-regressivo não-linear com baixo nível de ruído, o caos determinístico pode estar presente. Exemplos numéricos são dados

para séries do capítulo 1. O expoente de Lyapunov dominante, estimado a partir das redes neurais, é utilizado para complementar os resultados obtidos com os algoritmos de modelagem. Especial atenção é dedicada à determinação da dimensão de imersão a ser utilizada na reconstrução da dinâmica do sistema a partir da série temporal observada, importante para séries com pequeno comprimento. Testes estatísticos são empregados para avaliar os modelos de previsão obtidos com o algoritmo de modelagem.

No capítulo 5, testes estatísticos não-lineares são utilizados para identificar não-linearidades nos retornos das séries financeiras: IBOVESPA, BRASFUND, WHM e VALE. O algoritmo de modelagem no capítulo 4 é utilizado para investigar a existência de variáveis auto-regressivas nos sistemas dinâmicos geradores dessas séries e, em particular, se a existência de variáveis auto-regressivas não-lineares podem explicar as não-linearidades observadas. Como alternativa aos modelos estocásticos auto-regressivos não-lineares, são utilizados, também, modelos estocásticos não-lineares na variância (ARCH).

CAPÍTULO 1

SISTEMAS DETERMINÍSTICOS CAÓTICOS

Os sistemas determinísticos caóticos, empregados nos algoritmos deste texto, são apresentados a seguir. O teorema de Takens, que permite a reconstrução do sistema dinâmico a partir da série temporal observada e que desempenha papel central neste trabalho, é enunciado e exemplificado. Os atratores dos sistemas dinâmicos caóticos são conceituados e caracterizados através das suas propriedades geométricas.

1.1 CARACTERÍSTICAS DOS SISTEMAS DETERMINÍSTICOS CAÓTICOS

Sistemas determinísticos não-lineares podem apresentar comportamento semelhante ao de processos aleatórios. Estes sistemas são ditos caóticos (Formalmente, um sistema dinâmico determinístico é dito caótico se o expoente de Lyapunov dominante (def. em 2.2.2) é positivo (Brock, 1986, p.170)). Sistemas determinísticos caóticos apresentam grande variedade de caminhos no tempo (trajetórias, órbitas), com seu comportamento imitando processos com choques aleatórios, de forma diferente da classe de processos estocásticos lineares, onde somente quatro tipos de trajetórias podem ocorrer:

- 1) trajetórias oscilantes e estáveis (isto é, convergentes com oscilações de amplitude decrescente para o ponto de equilíbrio);
- 2) trajetórias oscilantes e explosivas (oscilações com amplitude crescente);
- 3) trajetórias não oscilantes e estáveis,
- 4) trajetórias não oscilantes e explosivas.

Baumol and Benhabib (1989, p.79) comentam que:

A lista de possíveis configurações 1) a 4), não é suficientemente ampla para propósitos econômicos, desde que as trajetórias de variáveis do mundo real são, frequentemente, mais complicadas. Oscilações podem não explodir ou serem amortecidas até desaparecerem.

Uma forma possível de obter variada classe de trajetórias é através dos sistemas determinísticos caóticos¹. Trajetórias caóticas podem ser caracterizadas por:

1) Apresentarem bruscas mudanças qualitativas em seu comportamento dinâmico, semelhantes a grandes choques aleatórios. Mas, de forma diferente dos processos submetidos a choques, os grandes movimentos da dinâmica são geradas internamente pelo sistema determinístico caótico.

2) Serem extremamente sensíveis a pequenas mudanças nos parâmetros do sistema e nas condições iniciais. Mudanças arbitrariamente pequenas nos parâmetros podem produzir trajetórias com características completamente diferentes. Trajetórias com condições iniciais arbitrariamente próximas divergem no tempo.

3) Nunca retornarem a ponto previamente alcançado, apesar de permanecerem evoluindo em região limitada.

Essas características motivaram grande número de trabalhos em séries temporais econômicas e financeiras, os quais procuram associar os grandes movimentos e o complexo equilíbrio de longo prazo de alguns dos sistemas dinâmicos geradores dessas séries, à presença de caos determinísticos. As técnicas empregadas para investigar caos nesses trabalhos são revistas no capítulo 2, com o auxílio dos sistemas determinísticos descritos a seguir.

1.2 SISTEMAS DETERMINÍSTICOS CAÓTICOS CLÁSSICOS

1.2.1 Tent

O sistema determinístico caótico tent(ou triangular) é definido pela equação de diferenças(Schuster,1988,p.26):

$$x_t = \Delta(x_{t-1}) = \begin{cases} 2x_{t-1}, & \text{se } x \leq 0.5, \\ 2(1 - x_{t-1}), & \text{se } x > 0.5. \end{cases} \quad (1.1)$$

Para quase todas as condições iniciais² $x_0 \in [0,1]$, as sequências de iteradas $\{x_t, t=1, 2, \dots\}$ dadas por:

$$x_1 = \Delta(x_0), x_2 = \Delta(x_1) = \Delta^2(x_0), \dots, x_n = \Delta^n(x_0), \dots$$

estão distribuídas de forma semelhante a um processo aleatório sobre o intervalo $[0,1]$. A figura 1.1.a mostra o gráfico do sistema tent e uma trajetória(órbita) gerada a partir da condição inicial x_0 .

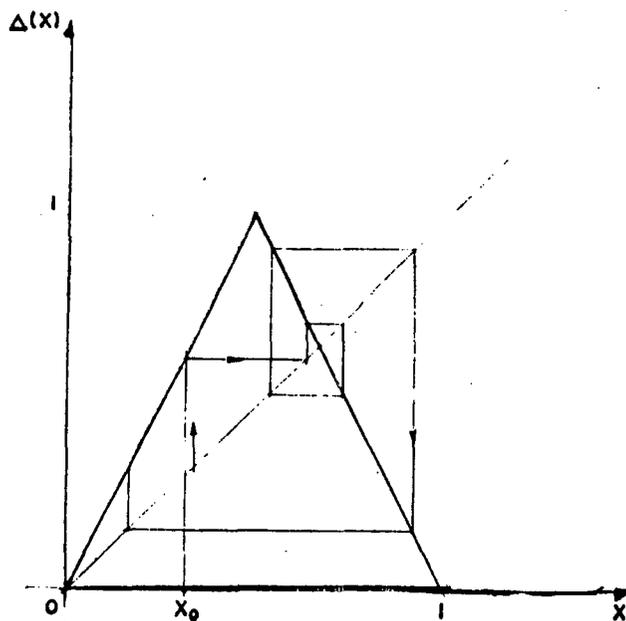


FIGURA 1.1.a Órbita(trajetória) do sistema determinístico tent (1.1) obtida a partir da condição inicial x_0 .

Apesar da sua aparente simplicidade, três importantes características dos sistemas determinísticos caóticos podem ser observadas nas sequências das iteradas(trajetórias, órbitas) do sistema dinâmico tent:

- 1) $\{x_t, t=1,2,\dots\}$ cobre de forma uniforme o intervalo $[0,1]$ (Lauwerier,1986,p.53);
- 2) erros de medida nas condições iniciais x_0 são ampliados exponencialmente(figura 1.1.b);

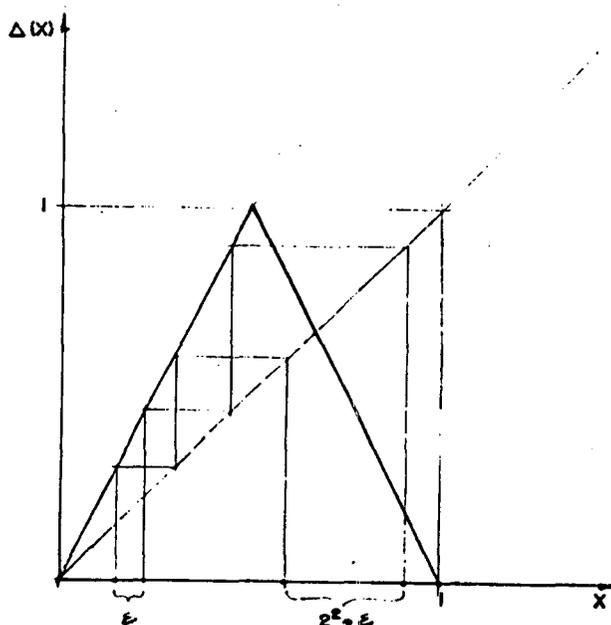


FIGURA 1.1.b Separação das trajetórias causada pela sensibilidade às condições iniciais do sistema dinâmico tent (1.1). As trajetórias separadas inicialmente por $\epsilon > 0$, no tempo $t=2$ estão separadas por $2^2 \epsilon$.

3) $\{x_t\}$ tem características de processo aleatório, apesar de ser processo determinístico, no sentido de que a função de autocorrelação e o espectro são os mesmos do ruído branco (Liu et al., 1992, p.526).

1.2.2 Logística

No sistema conhecido como Logística, a sequência $\{x_t\}$ é gerada pela equação de diferenças (Lauwerier, 1986, p.43):

$$x_t = \alpha x_{t-1}(1 - x_{t-1}), \quad (1.2)$$

com $x_0 \in [0,1]$ e $\alpha \in (0,4]$.

A Logística (1.2) é exemplo de sistema determinístico simples que exibe rica e surpreendente dinâmica (May, 1976). Lauwerier (1986, p.46) prova que:

Pontos de equilíbrio estáveis, ciclos limites e flutuações aparentemente aleatórias são geradas por essa equação de diferenças, em função do valor do parâmetro α . Para valores de $\alpha \in (0,3)$, o sistema é estável. Para $\alpha \in (0,1)$, as trajetórias convergem para o ponto fixo $x^* = 0$. Para $\alpha \in [1,3)$, as trajetórias convergem para um ponto fixo $x^* = 0$. Para $\alpha \in [3,4]$, as trajetórias podem ser periódicas ou caóticas. Entre 3 e $\alpha_{\infty} = 3.568\dots$ as trajetórias duplicam de período com α crescente. A partir de α_{∞} existem trajetórias periódicas e trajetórias caóticas e, também, trajetórias periódicas arbitrariamente próximas de trajetórias caóticas³.

¹ Muitos modelos na literatura de séries temporais procuram explorar não-linearidades, como forma de permitir trajetórias diferentes das dadas por modelos lineares. O modelo TAR (Threshold Auto-Regression, Tong and Lim, 1980) e o modelo GARCH (Generalized Autoregressive Conditional Heteroskedasticity, Bollerslev, 1986; 1987) são exemplos de modelos que podem produzir trajetórias diferentes das permitidas para modelos lineares. O primeiro pode produzir ciclos limites e o segundo trajetórias com concentrações de outliers.

² Todas as condições iniciais, com exceção de um conjunto de medida zero no sentido de Lebesgue, formam o conjunto de condições iniciais típicas.

Assim, pequenas mudanças no parâmetro α podem transformar uma trajetória caótica (periódica) em uma trajetória periódica (caótica). Esse fato ilustra a sensibilidade da dinâmica dos sistemas determinísticos caóticos aos parâmetros do modelo. Pequenas mudanças nos parâmetros podem gerar trajetórias com características totalmente diferentes.

1.2.3 Hénon

O sistema dinâmico (Hénon, 1976):

$$\begin{cases} x_t = 1 - a[x_{t-1}]^2 + y_{t-1} \\ y_t = bx_{t-1}, \end{cases} \quad (1.3)$$

é exemplo de sistema bidimensional que pode apresentar trajetórias caóticas. Para $a=1.4$ e $b=0.3$ (Hénon, 1976, p.72), quase todas as condições iniciais geram trajetórias que são atraídas para a região esboçada na figura 1.5.a. As trajetórias, ao evoluírem no tempo, percorrem as diferentes partes dessa região, sendo por ela aprisionadas. O conjunto de todos os pontos pertencentes a essa região é o atrator do sistema dinâmico caótico de Hénon.

1.2.4 A equação de Mackey-Glass discreta

A equação de diferenças (MacCaffrey et al., 1992, p.687):

$$x_{n+1} = ax_n + bx_{n-d}/(1 + [x_{n-d}]^k), \quad (1.4.1)$$

é a versão discreta da equação diferencial com defasagens

$$dx(t)/dt = ax(t) + bx(t-d)/(1+[x(t-d)]^k), \quad (1.4.2)$$

introduzida originalmente para modelar a produção e a perda de glóbulos brancos (Mackey and Glass, 1987). A figura 1.5.b apresenta um esboço do atrator para esse sistema dinâmico com os parâmetros $a=0.2$, $b=2.0$, $k=6$ e $d=2$ (MacCaffrey et al., 1992, p.687). Observa-se que de forma semelhante ao atrator de Hénon, o atrator da equação

³ A existência de trajetórias estáveis arbitrariamente próximas de trajetórias caóticas é explorada em uma nova fronteira dos caos: Controle de Sistemas Caóticos (Ott et al., 1990; Romeiras et al., 1992).

de Mackey-Glass discreta é uma região do R^2 com complexa geometria.

Os sistemas determinísticos caóticos (1.1) a (1.4) são descritos em termos de equações de diferenças. Existem outras formas de se obter sistemas determinísticos caóticos. Os sistemas caóticos de Ikeda e Lorenz são originados, respectivamente, de uma aplicação do plano complexo no plano complexo e de um sistema de três equações diferenciais ordinárias.

1.2.5 Ikeda

O sistema dinâmico de Ikeda (Casdagli, 1989, p.342), que modela a dinâmica de fenômenos ligados à emissão de raios laser, é uma aplicação do plano complexo no plano complexo definida por

$$F: R^2 \rightarrow R^2$$

$$(x, y) \rightarrow F(x, y) = (p + c_2(x \cos t - y \sin t), c_2(x \sin t + y \cos t)), \quad (1.5.1)$$

onde

$$t = c_1 - c_3 / (1 + x^2 + y^2). \quad (1.5.2)$$

O sistema dinâmico de Ikeda (1.5.1) e (1.5.2) tem neste trabalho os parâmetros $c_1=0.4$, $c_2=-0.6$ e $c_3=0.7$ e $p=1$ (Casdagli, 1989, p.342).

1.2.6 O modelo de Lorenz

O sistema de três equações diferenciais ordinárias (Lorenz, 1963):

$$\begin{cases} dx/dt = -\sigma(x + y) \\ dy/dt = -xz + rx - y \\ dz/dt = xy - bz, \end{cases} \quad (1.6)$$

com $(x, y, z) \in R^3$, $\sigma > 0$ e $b > 0$, é conhecido como modelo de Lorenz. Ele modela a convecção forçada de um líquido contra a ação da gravidade (por exemplo: formação de correntes de ar na atmosfera devido ao aquecimento do solo). Um dos primeiros comportamentos caóticos observados apresentou-se no modelo de Lorenz, no contexto da meteorologia. A sensibilidade do modelo às condições iniciais leva à impossibilidade de prever o comportamento do sistema para tempos distantes.

Os parâmetros para (1.6) neste trabalho são: $\sigma=10.0$, $r=0.28$ e $b=2.66$ (Koçak, 1986, p.138).

Os sistemas determinísticos caóticos (1.1) a (1.6) geram caos de "baixa dimensão", no sentido que a estrutura não-linear caótica pode ser detectada facilmente. Existem sistemas que geram caos de "alta dimensão". Caos de alta dimensão significa que a estrutura não-linear caótica é difícil de ser identificada, a partir de quantidade limitada de dados. Do ponto de vista operacional, não existe diferença entre caos de alta dimensão e processo aleatório (Casdagli, 1992, p.306) (4.3.1, capítulo 4). Um exemplo de sistema caótico de alta dimensão é dado pelo gerador de números pseudo aleatórios, apresentado a seguir.

1.2.7 Caos de alta dimensão

Hsieh (1991, p.1842) apresenta o seguinte exemplo de sistema caótico de alta dimensão:

Escolha-se um número A (por exemplo, 7^5) e um número primo P (por exemplo, $2^{32}-1$). Com um número inteiro z_0 , dito semente, entre 0 e P , geram-se novas sementes usando a regra

$$z_t = Az_{t-1} \pmod{P},$$

onde $x \pmod{y}$ significa o resto da divisão de x por y . Então, a sequência

$$\{x_t = z_t/P\},$$

está "distribuída uniformemente" no intervalo $[0,1]$.

Esse método de gerar números pseudo aleatórios, usado em muitos programas computacionais, é um processo determinístico caótico de alta dimensão, difícil de ser identificado pelos algoritmos de análise de sistemas dinâmicos caóticos (capítulo 2), considerados a partir de séries com comprimento limitado.

Outro exemplo de sistema que gera caos de alta dimensão é dado pela equação de Mackey-Glass (1.4.2) com os parâmetros $a=-0.1$, $b=0.2$, $d=100$ e $k=10$. Para investigar propriedades dinâmicas e geométricas desse sistema com algoritmos de análise de sistemas dinâmicos, são necessárias séries com comprimento superior a 10^5 (Casdagli, 1989, p.342).

1.2.8 Séries não-lineares com ruído

Além das séries determinísticas caóticas observadas nos sistemas determinísticos não-lineares (1.1) a (1.6), são utilizadas neste trabalho as séries não-lineares com ruído Hénon6, Hénon22, Lorenz1.5 e Lorenz12, obtidas, respectivamente, dos sistemas de Hénon e Lorenz adicionando-se ruído $\{\varepsilon_t\}$ iid. (independentes e identicamente distribuídos) com $\varepsilon_t \in N(0, \sigma^2)$.

O nível de ruído na série é definido, em porcentagem, pelo desvio padrão do ruído $\{\varepsilon_t\}$ dividido pelo desvio padrão da série determinística. A notação Hénon6, por exemplo, indica a série observada do sistema dinâmico de Hénon com nível de ruído de 6%.

1.3 GEOMETRIA DO CAOS

Os sistemas determinísticos caóticos podem ser caracterizados por suas propriedades geométricas. Por exemplo, o sistema dinâmico tent pode ser visto, intuitivamente, como a combinação dos mecanismos geométricos (figura 1.2.a):

- I) esticamento (stretching) uniforme do intervalo $[0,1]$ em duas vezes seu comprimento;
- II) dobramento (folding) pela metade do intervalo esticado.

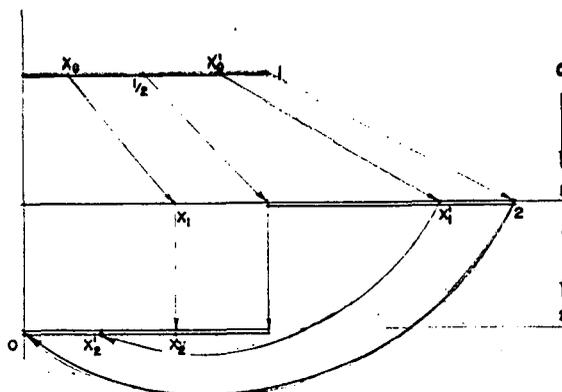


FIGURA 1.2.a Mecanismos geométricos do sistema dinâmico tent. O sistema evoluindo no tempo duplica o intervalo $[0,1]$ e, em seguida, dobra pela metade o intervalo duplicado. Esse processo de esticamento e dobra passa a imagem visual de que as trajetórias têm comportamento aleatório.

O mecanismo I) é responsável pela separação das condições iniciais arbitrariamente próximas, o que dá a sensibilidade às condições iniciais. O mecanismo II) obriga a órbita a ser limitada. A conjugação desses dois mecanismos faz a órbita andar "caoticamente" no intervalo $[0,1]$, passando a impressão visual de

um processo aleatório.

Os mecanismos de esticamento e dobra, observados no sistema *tent*, estão presentes em todos os sistemas dinâmicos caóticos caracterizados na seção 1.4 (atratores). Para esses sistemas, têm-se, de forma geral, a combinação dos mecanismos geométricos representados na figura 1.2.b.

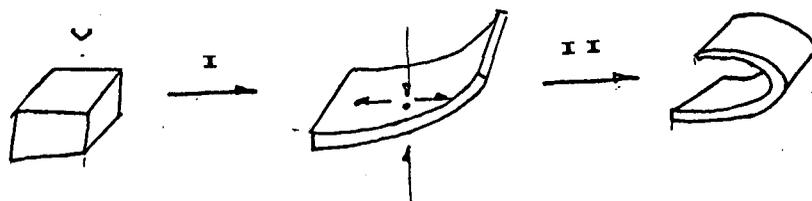


FIGURA 1.2.b Representação da ação dos mecanismos geométricos dos sistemas determinísticos caóticos sobre um elemento de volume arbitrário. Pontos inicialmente próximos são separados exponencialmente pelo mecanismo de esticamento (I). O mecanismo de dobra (II) os coloca novamente próximos. A conjugação desses dois mecanismos faz com que a trajetória tenha comportamento semelhante ao de um processo aleatório.

1.4 EXPOENTE DE LYAPUNOV

A separação de trajetórias com condições iniciais arbitrariamente próximas, causada pelos sistemas determinísticos caóticos e conhecida como sensibilidade às condições iniciais, pode ser quantificada através do expoente de Lyapunov. A figura 1.1.b mostra a separação de duas trajetórias do sistema dinâmico *tent*, obtidas a partir de duas condições iniciais x_0 e x_0^* separadas por ϵ . No tempo $t=2$ a distância entre $\Delta^2(x_0)$ e $\Delta^2(x_0^*)$ é $\epsilon 2^2$. Para $t=N$, tem-se que $\Delta^N(x_0)$ e $\Delta^N(x_0^*)$ estão separadas por (Schuster, 1988, p.28)

$$\epsilon 2^N,$$

que pode ser escrita na forma

$$\epsilon 2^N = \epsilon e^{N \log 2} = \epsilon e^{N \lambda},$$

com $\lambda = \log 2$. Assim, a taxa média pela qual a distância ϵ é multiplicada a cada iteração pelo sistema dinâmico *tent* é $e^{\log 2} > 1$. Essa taxa, maior do que um, mostra que as trajetórias obtidas a partir de condições iniciais x_0 e x_0^* , arbitrariamente próximas, divergem.

Para o sistema dinâmico unidimensional

$$x_{t+1} = f(x_t),$$

onde $f: \mathbb{R} \rightarrow \mathbb{R}$ é não-linear, tem-se que as trajetórias com condições iniciais x_0 e x_0^* , separadas por

$$\varepsilon = |x_0^* - x_0|$$

após N iterações ($t=1, \dots, N$; figura 1.3) estão separadas por

$$\varepsilon^* = |f^N(x_0^*) - f^N(x_0)|,$$

que pode ser escrita na forma

$$\varepsilon^* = |f^N(x_0^*) - f^N(x_0)| = \varepsilon e^{N\lambda},$$

onde e^λ é a taxa média de separação pela qual a distância ε é multiplicada a cada iteração do sistema dinâmico. Assim, se $e^\lambda > 1$, as trajetórias divergem com o tempo em evolução (as distâncias aumentam a cada iteração). Se $e^\lambda < 1$, as trajetórias convergem com o tempo em evolução (as distâncias diminuem a cada iteração). O valor de λ é chamado de expoente de Lyapunov do sistema dinâmico. Se $\lambda > 0$, então $e^\lambda > 1$ e as trajetórias divergem. Se $\lambda < 0$, então $e^\lambda < 1$ e as trajetórias convergem.

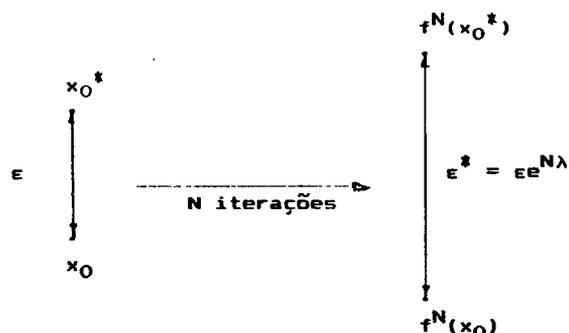


FIGURA 1.3 Separação das condições iniciais x_0 e x_0^* pelo sistema dinâmico $x_{t+1}=f(x_t)$. As condições iniciais separadas por ε após N iterações ($t=1, \dots, N$) estão separadas por ε^* , que pode ser escrita na forma $\varepsilon^* = \varepsilon e^{N\lambda}$, onde e^λ é a taxa média de separação. Se $e^\lambda > 1$, ε^* cresce com t e as trajetórias divergem. Se $e^\lambda < 1$, ε^* decresce com t e as trajetórias convergem.

O expoente de Lyapunov pode ser definido para os sistemas dinâmicos d -dimensionais com significado geométrico análogo ao dos sistemas dinâmicos unidimensionais (seção 2.2, capítulo 2). Sistemas dinâmicos determinísticos com expoente de Lyapunov positivo (sensibilidade às condições iniciais) são caracterizados

como caóticos (Eckmann et al., 1985; Nychka et al., 1992).

A figura 1.4 apresenta duas trajetórias para a Logística com $\alpha=3.568$, $t=1, \dots, 50$, e com as condições iniciais $x_0=0.3$ e $x_0^*=0.305$. A sensibilidade às condições iniciais (expoente de Lyapunov positivo; para a Logística $\lambda=0.69$) faz com que as trajetórias obtidas a partir de condições iniciais próximas separem-se com o tempo em evolução. Para $t < 12$, na escala utilizada, não é possível distinguir visualmente a diferença entre as duas trajetórias. Para $t \geq 12$, as trajetórias são muito diferentes, de tal forma que o conhecimento no tempo de uma delas não permite prever o comportamento da outra.

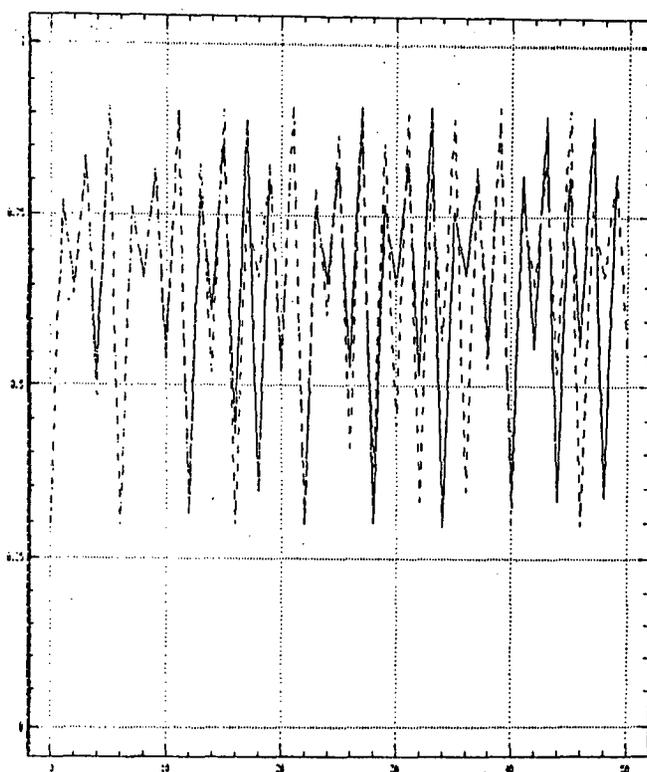


FIGURA 1.4 Trajetórias da Logística (1.2) com $\alpha=3.568$ obtidas das condições iniciais $x_0=0.3$ e $x_0^*=0.305$, $t=1, \dots, 50$. Para $t < 12$, na escala utilizada, não é possível distinguir visualmente a diferença entre as duas trajetórias. Para $t \geq 12$ as trajetórias são muito diferentes, de tal forma que o conhecimento no tempo de uma delas não permite prever o comportamento da outra.

Esse exemplo ilustra uma característica dos sistemas dinâmicos com expoente de Lyapunov positivo (sensibilidade às condições iniciais), que é a imprevisibilidade do sistema no longo prazo, causada por erros de medida nas condições iniciais.

1.5 PREVISÃO

Os mecanismos geométricos tornam a previsão das séries caóticas uma tarefa difícil a ser realizada com o instrumental tradicionalmente utilizado em séries temporais. Baumol and Benhabib(1989,p.79) observam que:

A previsão de séries temporais caóticas pode ser ineficiente com técnicas derivadas da extrapolação(em vários graus de sofisticação) e da estimação de modelos estruturais de previsão. As grandes flutuações geradas dentro dos sistemas dinâmicos caóticos tornam as técnicas que utilizam extrapolação pouco apropriadas. Modelos estruturais também são comprometidos, pois erros na estimação dos parâmetros do modelo podem levar a mudanças qualitativas na dinâmica do sistema caótico, tornando a previsão sem sentido.

Mas, se uma série temporal é observada de um sistema determinístico caótico, ela é previsível a curto prazo. O horizonte e a qualidade das previsões de curto prazo dependem dos mecanismos geométricos. Os modelos de previsão tradicionalmente empregados em Séries Temporais devem ser substituídos por modelos de previsão não-lineares(capítulo 3) apropriados a esse tipo de série.

1.6 ATRADORES

1.6.1 Conceito

Sistemas dinâmicos que contraem volumes no espaço de fase são ditos dissipativos. O volume da região do espaço de fase, contendo todas as condições iniciais típicas, é contraído pelo sistema dinâmico dando origem a uma região com volume zero, chamada de atrator do sistema dinâmico(Schuster,1988,p.103).

Nos sistemas dinâmicos(dissipativos) caóticos, apesar da contração de volume, comprimentos não são contraídos em todas as direções e pontos inicialmente próximos tornam-se exponencialmente separados com o tempo em evolução, até que o processo de dobra os coloque novamente próximos(figura 1.2.b). Esse mecanismo de esticamento e dobra de regiões faz com que a geometria do atrator seja complexa, dando origem ao que se chama de atrator estranho.

Sistemas dinâmicos com atratores estranhos são

caracterizados neste trabalho como caóticos⁴(Grassberger and Procaccia,1983a,b).

Nas figuras 1.5.a e 1.5.b estão os esboços dos atratores dos sistemas dinâmicos de Hénon e Mackey-Glass. A partir de condições iniciais típicas, as trajetórias no tempo, para esses dois sistemas dinâmicos, percorrem as diferentes partes dos respectivos atratores, sem nunca encontrar pontos passados da trajetória.

1.6.2 Reconstrução do atrator e da dinâmica de um sistema

O importante resultado apresentado a seguir, devido a Packard et al.(1980) e colocado em bases matemáticas sólidas por Takens(Casdagli,1986,p.171), permite a aplicação de técnicas da análise de sistemas dinâmicos não-lineares em séries temporais univariadas. A partir da observação de apenas uma das variáveis do sistema dinâmico, é possível reconstruir a dinâmica e o atrator do sistema no espaço de fase. A reconstrução dada pelo teorema de Takens consiste em obter um sistema de equações de diferenças, envolvendo apenas uma das variáveis do sistema original e suas defasagens, que têm as mesmas propriedades geométricas e dinâmicas do sistema dinâmico d-dimensional do qual se observa a variável, utilizada no processo de reconstrução.

O teorema de Takens

O teorema de Takens, como enunciado em Casdagli(1992,p.305) para variáveis observadas em tempo discreto, é mais apropriado para os objetivos deste trabalho, cuja ênfase é em séries temporais:

Supondo-se que uma série univariada $\{x_n, n=1, \dots\}$ é observada do sistema determinístico d-dimensional:

$$s(t) = F^t(s(0)), \quad (1.7)$$

⁴ Não existe definição formal de atrator estranho, aceita amplamente. Grebogi et al.(1984) argumentam que estranho é diferente de caótico, mostrando a existência de atrator estranho em sistema dinâmico que não é caótico. Estranho refere-se, segundo Grebogi et al., à geometria do atrator e caótico refere-se ao comportamento da dinâmica dado por sistemas com sensibilidade às condições iniciais.

com

$$x_n = h(s(n\tau_s)), \quad (1.8)$$

onde $s(t) \in \mathbb{R}^d$ são os estados do sistema no tempo t , a aplicação $F: \mathbb{R}^d \rightarrow \mathbb{R}^d$ dá a evolução da dinâmica do sistema no tempo a partir da condição inicial $s(0)$, τ_s é a taxa de amostragem (periodicidade de amostragem) e $h: \mathbb{R}^d \rightarrow \mathbb{R}$ é o observador.

Takens prova que:

1) Para $m > 2d$ e para a defasagem τ , as propriedades métricas e dinâmicas do sistema com estados $s(t) \in \mathbb{R}^d$ podem ser obtidas da sequência de vetores $\{X_n, n=1, \dots\}$, onde

$$X_n = (x_n, x_{n-\tau}, \dots, x_{n-(m-1)\tau}) \in \mathbb{R}^m. \quad (1.9)$$

2) Existe f não-linear diferenciável

$$f: \mathbb{R}^m \rightarrow \mathbb{R}^m, \quad (1.10)$$

tal que para todos os inteiros p e τ ,

$$x_{n+p} = f^p(X_n) = f^p(x_n, x_{n-\tau}, \dots, x_{n-(m-1)\tau}). \quad (1.11)$$

onde $f^p = f \circ f \circ \dots \circ f$, é a "p-ésima iterada de f ."

O símbolo " \circ " acima indica a composição de funções. O valor de m em (1.9) e (1.11) é dito dimensão de imersão. O método de reconstrução, dado pelo teorema de Takens, é conhecido como MÉTODO DAS DEFASAGENS. A escolha dos valores de m e τ utilizados no processo de reconstrução é discutida na próxima seção.

No apêndice B procura-se, através de exemplos simples, ilustrar o significado do Teorema de Takens para sistemas dinâmicos.

Reconstrução

O teorema de Takens mostra que é possível reconstruir a dinâmica e o atrator de um sistema d -dimensional não conhecido (algebricamente), do qual observa-se apenas uma variável. Dada a série temporal $\{x_n, n=1, \dots, T\}$, escolha-se uma dimensão de imersão m , uma defasagem τ e construa-se a sequência de vetores m -dimensionais

$$X_n = (x_n, x_{n-\tau}, \dots, x_{n-(m-1)\tau}) \in \mathbb{R}^m,$$

$n=1, \dots, T$. As propriedades métricas e dinâmicas do sistema dinâmico d -dimensional, do qual observa-se apenas a variável x_n , podem ser obtidas da sequência de vetores $\{X_n, n=1, \dots, T\}$.

A qualidade da reconstrução dada pelo método das defasagens, para séries temporais com comprimento finito, depende de τ e de m . Neste trabalho, τ será número inteiro positivo⁵. O valor $m > 2d$ é cota superior para o processo de reconstrução pelo método das defasagens, reconstruindo globalmente o sistema dinâmico no espaço m -dimensional. Nas aplicações em geral, basta que as reconstruções sejam locais. Para reconstrução local existe $m_{\min} < 2d$ ⁶, valor este que possibilita uma boa reconstrução da dinâmica e do atrator no espaço de fase. A reconstrução local, que pode ser feita em dimensão $m < 2d$, exige séries com menor comprimento.

É importante observar que a maioria dos trabalhos que investigam caos em séries temporais econômicas e financeiras⁷ e, também, em ciências naturais, procuram mostrar a existência de um atrator estranho, empregando a primeira parte do Teorema de Takens, a partir de técnicas que ignoram a existência de uma dimensão mínima para a reconstrução do sistema pelo método das defasagens. Para séries do mundo real, em geral com pequeno comprimento⁸, esses procedimentos que ignoram a dimensão de imersão ideal não são recomendados, pois o número de dados necessários a essas técnicas cresce rapidamente com a dimensão de imersão. Além disso, ignoram a existência da função f em (1.11).

Neste trabalho, pretende-se investigar modelos de previsão de séries determinísticas caóticas, utilizados na Física a partir de 1987 (Farmer and Sidorowich, 1987) e implementados aqui com técnicas de regressão não-paramétrica disponíveis na literatura econométrica (capítulo 3). Esses modelos procuram obter aproximações da função f dada pela segunda parte do teorema de

⁵ O valor de τ pode ser determinado através da função de autocorrelação (Abarbanel et al., 1990).

⁶ Para uma justificativa matemática sobre a utilização de dimensão de imersão menor que a exigida pelo teorema de Takens, ver Ott et al. (1990).

⁷ A investigação da existência de atratores em séries econômicas e financeiras é feita em Brock (1986), Brock and Sayers (1988), Scheinkman and LeBaron (1989), Frank and Stengos (1988a, b) e Ramsey et al. (1990). Scheinkman and LeBaron, com dados do Center for Research in Security Prices (CRSP), indicam a existência de atrator caótico para o sistema dinâmico associado a essa série financeira.

Takens, que possibilitam fazer a previsão x_{t+p} , p passos à frente, de x_t . Pretende-se, também, abordar o problema de determinar a dimensão de imersão mínima(ou ideal), importante para séries temporais do mundo real e que os pesquisadores de caos em Física afirmam existir(quando relaxam a hipótese que $m > 2d$), mas à qual não dedicam atenção especial, pois trabalham com séries com grande comprimento.

1.6.3 Exemplos de reconstrução

O atrator de Hénon

Um exemplo bem conhecido da reconstrução do atrator de um sistema dinâmico, a partir da observação de apenas uma variável, é dado pelo sistema determinístico caótico de Hénon (1.3) com $a = 1.4$ e $b = 0.3$.

Escolha-se a dimensão de imersão $m=2$ (menor que a exigida pelo teorema de Takens), a defasagem $\tau=1$ e considere-se apenas a variável x_n como observada. A sequência de vetores

$$\{X_n = (x_n, x_{n-1}), n=1, \dots, T\} \in \mathbb{R}^2,$$

reconstrói o atrator do sistema dinâmico. Através dessa sequência pode-se investigar as propriedades métricas e dinâmicas do sistema original. Na figura 1.5.a, estão plotados 1500 vetores da sequência bidimensional X_n . Observa-se na figura o atrator de Hénon, reconstruído no espaço de fase, ocupando região do espaço \mathbb{R}^2 . Para quase todas as condições iniciais $X_0 = (x_0, y_0)$, as trajetórias do sistema de Hénon são atraídas e aprisionadas por essa região. Na reconstrução, foi utilizada a coordenada x_n . Resultado análogo é obtido, segundo o teorema de Takens, com a coordenada y_n .

⁸ Os termos séries com pequeno comprimento e grande comprimento são empregados em vários lugares neste texto. O número de observações de uma série com pequeno comprimento ou com grande comprimento depende do contexto. Por exemplo, para investigar a existência de atrator, uma série com 10000 observações pode ser com pequeno comprimento, enquanto para investigar a sensibilidade às condições iniciais, pode ser com grande comprimento. Para aplicações em economia e finanças aceita-se a colocação de Liu et al.(1992,p.S29): " 500(5900) observações é uma série com comprimento grande mas plausível em Macroeconomia(Finanças), o que corresponde a aproximadamente 40 anos de dados mensais(20 anos de dados diários)".

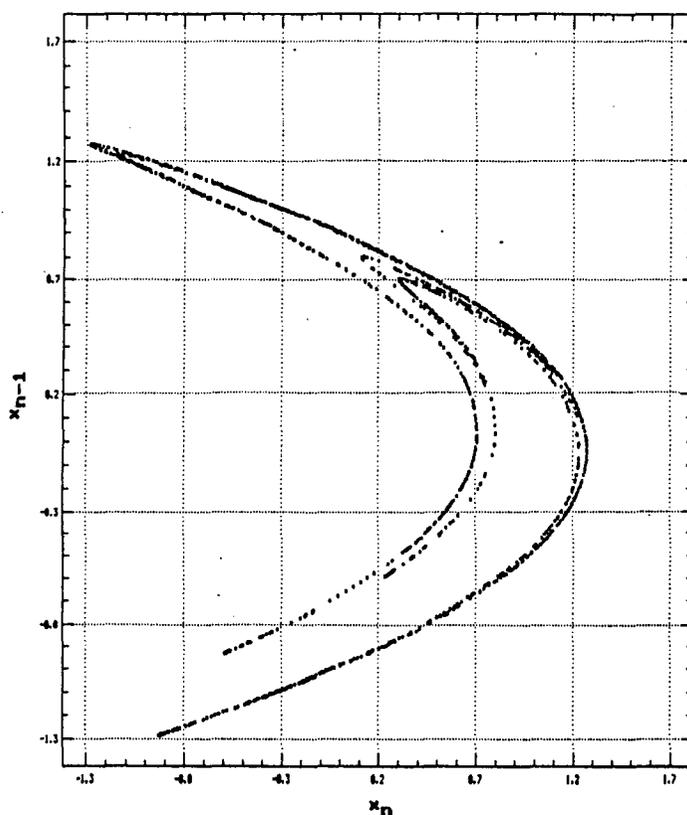


FIGURA 1.5.a Atrator caótico do sistema dinâmico de Hénon(1.3), obtido da sequência $\{X_n=(x_n, x_{n-1}), n=1, \dots, 1500\}$, através do teorema de Takens(Método das defasagens com $m=2$).

O atrator de Mackey-Glass

Dada a série $\{x_n, n=1, \dots, 1500\}$, gerada pela equação de Mackey-Glass

$$x_{n+1} = 0.2x_n + 2x_{n-2}/(1+[x_{n-2}]^6),$$

escolha-se $m=2$ e $\tau=1$, obtendo-se a sequência bidimensional de vetores

$$\{X_n=(x_n, x_{n-1}), n=1, \dots, 1500\} \in \mathbb{R}^2,$$

que reconstrói o sistema dinâmico. Na figura 1.5.b, esses vetores estão plotados. A região esboçada nessa figura é o atrator do sistema dinâmico associado a essa equação de diferenças.

Observa-se, nas figuras 1.5.a e 1.5.b, que os atratores(caóticos) desses dois sistemas são regiões do espaço \mathbb{R}^2 com complicada geometria e com distribuição de pontos em geral não uniforme. Os conceitos usuais de dimensão de objetos geométricos não são aplicáveis a essas regiões. No capítulo 2, os atratores são caracterizados a partir de conceitos apropriados de dimensão, que levam em consideração aspectos geométricos e probabilísticos. A presença de atrator com dimensão não

inteira(fractal) em sistema dinâmico determinístico é traço característico da presença de caos no sistema(Grassberger and Procaccia,1983a,b; Liu et al.,1992).

Atratores com dimensão não-inteira são chamados de atratores estranhos(Grassberger and Procaccia 1983a).

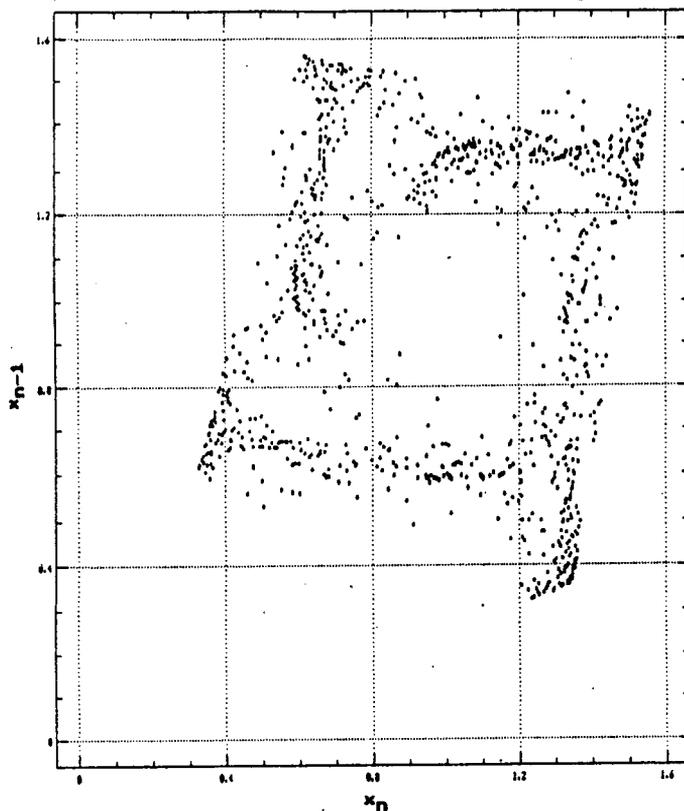


FIGURA 1.5.b Atrator caótico do sistema dinâmico de Mackey-Glass (1.4), obtido da sequência $\{X_n = (x_n, x_{n-1}), n=1, \dots, 1500\}$, através do teorema de Takens(Método das defasagens com $m=2$).

CAPÍTULO 2

IDENTIFICAÇÃO DE SISTEMAS DETERMINÍSTICOS CAÓTICOS

Métodos clássicos

A identificação de sistemas determinísticos caóticos é feita, geralmente, através do cálculo da dimensão do atrator e do expoente de Lyapunov. Dimensão não-inteira caracteriza o atrator como caótico. Expoente de Lyapunov positivo indica sensibilidade às condições iniciais, que é uma das características dos sistemas determinísticos caóticos. Neste capítulo são apresentadas definições de dimensão e do expoente de Lyapunov, algoritmos para obter esses invariantes geométricos e dinâmicos e resultados computacionais. As limitações desse tipo de análise na identificação de sistemas determinísticos caóticos também são abordadas.

2.1 DIMENSÃO

Existem várias definições de dimensão (Grassberger and Procaccia, 1983a,b; Farmer, 1982; Farmer et al., 1983), nem todas dando o mesmo valor para um dado objeto. As mais importantes podem ser englobadas em dois tipos:

- a) dimensões que dependem somente das propriedades métricas;
- b) dimensões que dependem da frequência com a qual uma órbita típica visita as diferentes partes do atrator.

A capacidade de um conjunto e a dimensão de correlação são definidas a seguir. A primeira depende somente das propriedades métricas enquanto que a segunda leva em consideração a distribuição dos pontos no atrator.

2.1.1 Capacidade

A capacidade de um conjunto $A \subset \mathbb{R}^D$, A limitado (Farmer et al., 1983, p.157) é dada por

$$d_c = \lim_{\epsilon \rightarrow 0} \frac{\log(M(\epsilon))}{\log(1/\epsilon)}, \quad (2.1)$$

onde $M(\epsilon)$ é o número mínimo de cubos de dimensão d e aresta ϵ necessário para recobrir o conjunto A .

A capacidade d_C para os conjuntos dados por pontos isolados, linhas e áreas, coincide com a dimensão utilizada na geometria elementar. Assim, a definição de capacidade amplia o conceito usual da dimensão de objetos geométricos. Para um ponto no plano (figura 2.1.a), tem-se $M(\epsilon)=1$ para todo ϵ . Então,

$$d_C = \lim_{\epsilon \rightarrow 0} \frac{\log M(\epsilon)}{\log(1/\epsilon)} = \lim_{\epsilon \rightarrow 0} \frac{\log 1}{\log(1/\epsilon)} = 0.$$

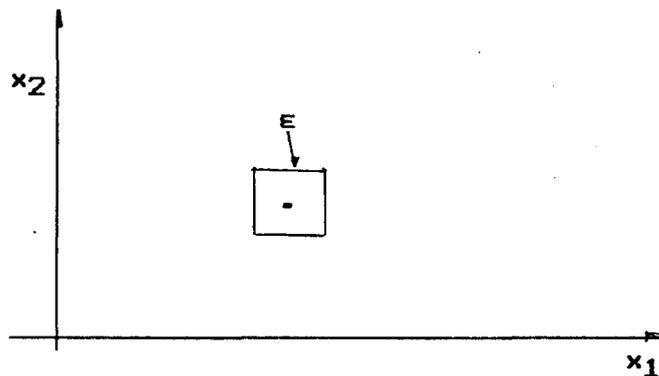


FIGURA 2.1.a Recobrimento de um ponto isolado por cubos de dimensão 2 e aresta ϵ .

Para uma curva plana C de comprimento l (figura 2.1.b), tem-se $M(\epsilon) \propto L/\epsilon$, isto é, o número de cubos necessários para fazer o recobrimento da curva é proporcional ao comprimento da curva. Então,

$$d_C = \lim_{\epsilon \rightarrow 0} \frac{\log M(\epsilon)}{\log(1/\epsilon)} = \lim_{\epsilon \rightarrow 0} \frac{\log(L/\epsilon)}{\log(1/\epsilon)} =$$

$$\lim_{\epsilon \rightarrow 0} \frac{\log L - \log \epsilon}{-\log \epsilon} = \lim_{\epsilon \rightarrow 0} \left(1 - \frac{\log L}{\log \epsilon} \right) = 1,$$

pois $\log \epsilon \rightarrow -\infty$, com $\epsilon \rightarrow 0$.

De forma análoga, para uma superfície plana com área A , tem-se $M(\epsilon) \propto A/\epsilon^2$. Então,

$$d_C = \lim_{\epsilon \rightarrow 0} \frac{\log M(\epsilon)}{\log(1/\epsilon)} = \lim_{\epsilon \rightarrow 0} \frac{\log(A/\epsilon^2)}{\log(1/\epsilon)} = 2.$$

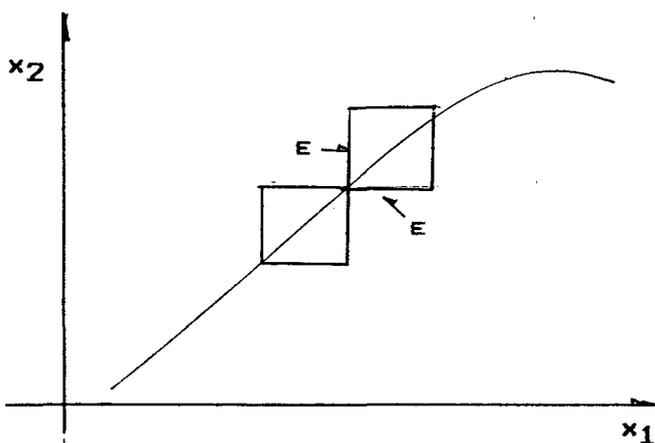
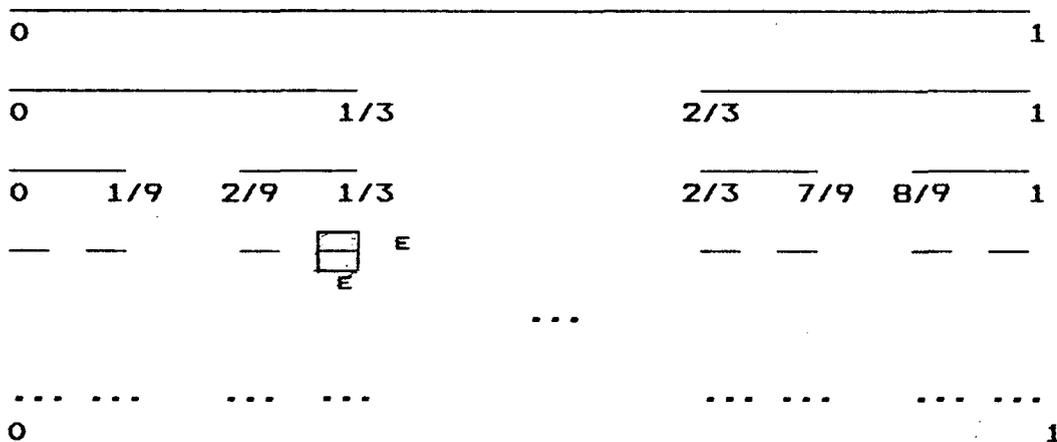


FIGURA 2.1.b Recobrimento de uma curva C de comprimento 1 por cubos de dimensão 2 e aresta ϵ .

Para conjuntos mais gerais, chamados de fractais (Mandelbrot, 1982), d_C pode ser não-inteira. A construção a seguir mostra um objeto geométrico com dimensão não-inteira.

Dado o segmento de reta com comprimento 1, retire-se desse segmento o terço médio. Em seguida retirem-se, sucessivamente, os terços médios dos segmentos resultantes. As etapas desse processo de construção estão esquematizadas abaixo.



O número $M(\epsilon)$ de cubos de dimensão 2 e aresta $\epsilon=1/3^n$ necessários para recobrir o conjunto resultante, após a retirada dos terços médios, na n -ésima etapa desse procedimento é:

$$M(\epsilon) = 2^n.$$

Então,

$$d_C = \lim_{\epsilon \rightarrow 0} \frac{\log M(\epsilon)}{\log(1/\epsilon)} = \lim_{n \rightarrow \infty} \frac{\log 2^n}{\log 3^n} =$$

$$\lim_{n \rightarrow \infty} \frac{\ln 2}{\ln 3} \approx 0.631.$$

O conjunto obtido com o procedimento de retirar o terço médio é conhecido como conjunto de Cantor. Esse conjunto tem dimensão não-inteira (capacidade $d_C = 0.631$), isto é, o conjunto de Cantor é um objeto fractal.

A existência de atrator com dimensão não-inteira (fractal) caracteriza neste trabalho o sistema dinâmico como caótico (Liu et al., 1992), embora existam sistemas dinâmicos não-caóticos com atratores de dimensão não-inteira (Grebogi et al., 1984; nota 4, seção 1.6). Atratores com dimensão fractal são chamados de atratores estranhos (Grassberger and Procaccia, 1983a). A figura 2.1.c apresenta o atrator caótico (estranho) do sistema dinâmico de Hénon, reconstruído pelo método das defasagens com série de comprimento $T=1500$ e na dimensão de imersão m igual a 2. A valor da dimensão (capacidade) aceita para esse atrator estranho é $d_C \approx 1.26$ (Grassberger and Procaccia, 1983b).

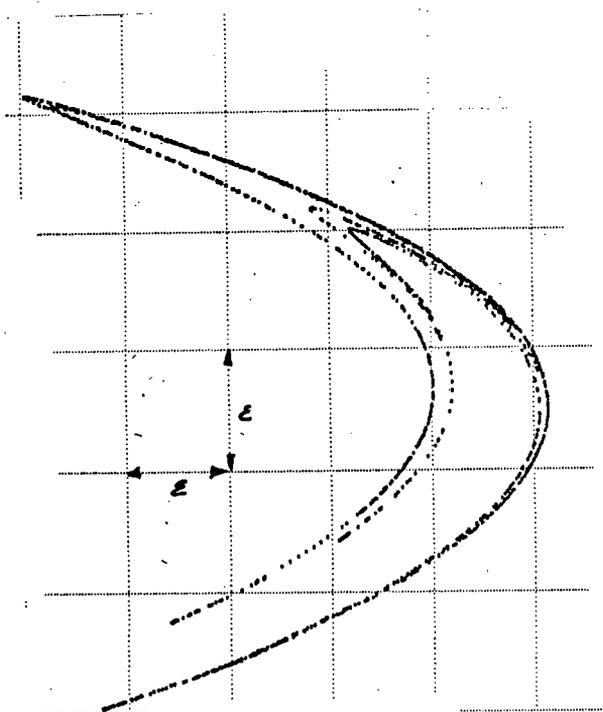


FIGURA 2.1.c Recobrimento do atrator do sistema dinâmico de Hénon (1.3) por cubos de dimensão 2 e aresta ϵ . A dimensão (capacidade) aceita para esse atrator estranho é $d_C \approx 1.26$.

Na definição de capacidade, os cubos usados no recobrimento do atrator entram com o mesmo peso na determinação de d_C , independentemente dos números de pontos por eles contidos. Assim, a capacidade não é sensível à frequência com a qual a trajetória

visita as diferentes partes do atrator.

O cálculo de d_c pode ser difícil ou até impraticável, inclusive para atratores de sistemas dinâmicos caóticos com baixa dimensão (Ben-Mizrachi et al., 1984, p.975). Na referência Grassberger (1983a, p.348), para estimar d_c para o sistema dinâmico de Hénon são necessários 200000 pontos para assegurar a existência do limite em (2.1). Para o modelo de Lorenz, com o mesmo número de pontos, não é obtida, nessa referência, a convergência do limite em (2.1).

Esses exemplos mostram que o cálculo da capacidade de atratores exige grande número de pontos e grande capacidade computacional. A dimensão definida a seguir evita parcialmente os problemas apontados na determinação experimental da capacidade d_c .

2.1.2 Dimensão de correlação

A dimensão de correlação é definida em Grassberger and Procaccia (1983a, p.347):

Considere-se a série temporal $\{X_t \in \mathbb{R}^n, t=1, \dots, T\}$ no atrator caótico. Devido a sensibilidade às condições iniciais as trajetórias divergem exponencialmente, implicando que (X_i, X_j) , $i \neq j$, são dinamicamente não correlacionados, como se fossem pontos de uma sequência gerada aleatoriamente. Mas, tais pontos estão imersos no atrator, implicando que são espacialmente correlacionados. A correlação espacial pode ser medida pela integral de correlação

$$C(\epsilon) = \lim_{T \rightarrow \infty} (1/T^2) \{ \#(i,j) : \|X_i - X_j\| < \epsilon \} \quad (2.2)$$

onde $\| \cdot \|$ é alguma norma no \mathbb{R}^n e $\#(i,j)$ indica o número de pontos com a propriedade que especifica o conjunto.

A integral de correlação mede a fração de pontos do atrator que estão a uma distância menor do que ϵ uns dos outros.

Se para $\epsilon \rightarrow 0$, a integral de correlação $C(\epsilon)$ segue a lei de escala

$$C(\epsilon) \propto \epsilon^V, \quad (2.3)$$

o valor de ν é chamado de expoente de correlação ou dimensão de correlação (Grassberger and Procaccia, 1983a, p.347).

A dimensão de correlação é cota inferior para capacidade. Vale, em geral, a relação $\nu \leq d_c$. A igualdade é realizada quando a distribuição dos pontos no atrator é uniforme (Grassberger and Procaccia, 1983b, p.197).

Para a Logística, com $\alpha=3.5699456$, tem-se $\nu=0.500 < 0.538=d_c$, mostrando que a distribuição dos pontos no atrator da Logística não é uniforme, enquanto que para o atrator do modelo de Lorenz, $\nu=d_c=2.06$ mostra que a distribuição dos pontos nesse atrator é uniforme (Grassberger and Procaccia, 1983a, p.348).

Para atratores de sistemas determinísticos caóticos com baixa dimensão(?), as diferenças numéricas entre as várias definições de dimensão são geralmente pequenas, mas existem exemplos de atratores onde elas são muito grandes (Smith, 1992, p.330).

A dimensão de correlação é a dimensão mais empregada na identificação de caos em sistemas dinâmicos não-lineares. Duas razões podem ser apontadas para justificar sua ampla utilização:

1) das definições de dimensão, é a que exige menor capacidade computacional e que tem algoritmos com melhores propriedades de convergência;

2) testes estatísticos para detectar não-linearidades em séries temporais são construídos a partir da integral de correlação.

2.1.3 Algoritmos

A definição alternativa da integral de correlação em Hsieh (1989, p.342), apresentada a seguir, facilita as estimativas da dimensão de correlação:

A integral de correlação pode, também, ser definida por

$$C(\epsilon) = \lim_{T \rightarrow \infty} \frac{2}{T(T-1)} \sum_{1 \leq i < j \leq T} I_{\epsilon}(Z_i, Z_j), \quad (2.4)$$

onde $\{Z_t, t=1, \dots, T\}$ é uma série temporal de vetores de dimensão d e I_{ϵ} é a função indicadora definida por

$$I_{\epsilon} = \begin{cases} 1, & \text{se } \|Z_i - Z_j\| < \epsilon \\ 0, & \text{se } \|Z_i - Z_j\| \geq \epsilon, \end{cases}$$

e $\| \cdot \|$ indica alguma norma no espaço \mathbb{R}^d .

Para T finito e para dado valor de ϵ , estima-se a integral de correlação $C(\epsilon)$ por

$$C(T, \epsilon) = \frac{2}{T(T-1)} \sum_{1 \leq i < j \leq T} I_{\epsilon}(Z_i, Z_j). \quad (2.5)$$

A dimensão de correlação $\nu(T)$ é estimada de (2.3), através de

$$\log C(T, \epsilon) \simeq k + \nu \log \epsilon, \quad (2.6)$$

por regressão linear, na parte linear do gráfico de $\log C(T, \epsilon)$ vs. $\log \epsilon$, como é recomendado por Grassberger and Proccacia (1983a, b).

Outra opção, segundo Brock and Baeck (1991, p.699), é estimar ν por

$$\nu(T, \epsilon, \epsilon + \Delta \epsilon) = (\log C(T, \epsilon + \Delta \epsilon) - \log C(T, \epsilon)) / (\log(\epsilon + \Delta \epsilon) - \log \epsilon), \quad (2.7)$$

que é uma estimativa de

$$\nu = d \log C(T, \epsilon) / d \log \epsilon.$$

A integral de correlação pode ser estimada para processos univariados a partir da reconstrução do atrator, pelo método das defasagens. Seja $\{x_t, t=1, \dots, T\}$ uma série univariada. Para dada dimensão de imersão m , construa-se a sequência de vetores de dimensão m (chamados de m -histórias):

$$Z_t = x_t^m = (x_t, x_{t-1}, \dots, x_{t-(m-1)}),$$

com $t=m, \dots, T$.

A integral de correlação $C(T, \epsilon, m)$, para essa dimensão de imersão, é estimada por

$$C(T, \epsilon, m) = \frac{2}{T(T-1)} \sum_{m \leq i < j \leq T} I_{\epsilon}(x_i^m, x_j^m), \quad (2.5)$$

e a dimensão de correlação $\nu(T, m)$ é estimada de (2.6) ou (2.7), com $C(T, \epsilon, m)$ no lugar de $C(T, \epsilon)$.

As estimativas da dimensão de correlação dependem dos valores de ϵ considerados. Em (2.6), escolhe-se visualmente a parte linear do gráfico $\log C(T, \epsilon, m)$ vs $\log \epsilon$ e estima-se ν através de regressão linear. Em (2.7), um procedimento comum é transformar a série para o intervalo $[0, 1]$ e escolher ϵ na forma $\epsilon = 0.9^j$, $j = 1, \dots, k$, (Brock and Baek, 1991; Liu et al, 1992, p.529). Outra possibilidade é escolher valores de ϵ expressos em termos do desvio padrão da série (DP). Os valores de ϵ são, geralmente, $\epsilon = 0.50, 0.75, 1.00, 1.25$ e 1.50 DP (Hsieh, 1989, p.348). Neste trabalho, as estimativas da dimensão de correlação são feitas através de (2.6) ou (2.7) com $\epsilon = 1/2^i$, $i = 1, 2, \dots$. Também são utilizados valores de ϵ como em Hsieh (1989).

2.1.4 Distinguindo sistemas determinísticos caóticos de ruído

A dimensão de correlação permite distinguir sistema determinístico caótico de um processo independente e identicamente distribuído (iid.). Grassberger and Procaccia (1983a, b) mostram que:

Imergindo-se o sistema dinâmico no espaço m -dimensional e fazendo m crescer, então, se $\{x_t\}$ é iid., $\nu(m) = m$ para cada m . No caso de a dinâmica ser governada por sistema determinístico caótico, $\nu(m)$ não cresce indefinidamente com m , atingindo valor ν^* , a partir do qual se torna independente de m .

Para um sistema determinístico caótico com baixo nível de ruído (definição em 1.2.8), o ruído não destrói a estrutura caótica (fractal) do atrator. Ben-Mizrachi et al. (1984, p.976) mostram que:

Imergindo-se o sistema dinâmico em espaço m -dimensional tem-se:

1) em escalas maiores que a do nível de ruído, $C(\epsilon) \propto \epsilon^m$,

donde $\nu=m$;

2) em escalas menores que a do nível de ruído, $C(\epsilon) \propto \epsilon^\nu$,
donde ν é a dimensão de correlação.

O gráfico da curva $\log C(\epsilon)$ vs. $\log \epsilon$ dada por (2.6) apresenta mudança da declividade (inclinação, coeficiente angular) no ponto em que ϵ iguala o nível de ruído, dando um critério para caracterizar atratores com ruído. O valor de ϵ onde ocorre a mudança da declividade do gráfico informa também sobre o nível de ruído no sistema, apontando as escalas que dizem respeito ao comportamento determinístico caótico e ao ruído.

2.1.5 Resultados numéricos

Foram estimadas, com séries de comprimento $T=1500$, as dimensões de correlação para os sistemas determinísticos caóticos Hénon ($a=0.3$, $b=1.4$) e Mackey-Glass ($a=0.2$, $b=2$, $k=6$, $d=2$), para o processo estocástico ARCH(2) $\{\epsilon_t\}$ dado por

$$z_t = \epsilon_t / h_t^{1/2}$$

onde $z_t \in N(0,1)$ iid. e

$$h_t = \alpha_0 + \alpha_1 \epsilon_{t-1}^2 + \alpha_2 \epsilon_{t-2}^2,$$

com $\alpha_0=0.05$, $\alpha_1=0.5$ e $\alpha_2=0.2$ e para o processo iid. normal $N(0,1)$.

Nas figuras 2.4.a e 2.2.a estão, respectivamente, os gráficos de $\log C(T,m,\epsilon)$ vs. $\log \epsilon$ para os sistemas dinâmicos de Hénon e Mackey-Glass com $m=2, \dots, 5$. Para cada m existe uma "parte linear" no gráfico correspondente. A declividade da reta, ajustada na parte linear do gráfico, é a dimensão de correlação para o sistema dinâmico na dimensão m , $m=2, \dots, 5$.

Para o sistema de Hénon na dimensão de imersão $m=2$, da regressão linear

$$\log C(1500, \epsilon, 2) = \underset{(0.03)}{-0.27} + \underset{(0.01)}{1.19} \log \epsilon,$$

$$R^2 = 0.98,$$

obtem-se, como estimativa para a dimensão de correlação, $\nu = \nu(1500, 2) = 1.19$, com os erros padrões das estimativas dos coeficientes da regressão linear dados entre parênteses. O valor de R^2 é o coeficiente de determinação da regressão.

Para a equação de Mackey-Glass na dimensão de imersão $m=3$, tem-se

$$\log C(1500, \epsilon, 3) = 0.06 + 1.65 \log \epsilon,$$

$$(0.13) \quad (0.02)$$

$$R^2 = 0.98.$$

A dimensão de correlação estimada é, então, $\nu = \nu(1500, 3) = 1.65$.

Observa-se que, nas figuras 2.4.a e 2.2.a, as partes lineares dos gráficos para $m=2, \dots, 5$, são aproximadamente paralelas, isto é, a declividade ν é independente de m . Isso permite concluir, através do algoritmo de Brassberger-Procaccia (seção 2.1.4), que os dois sistemas são caóticos, com as dimensões de correlação estimadas acima.

Nas figuras 2.6.a e 2.6.b, estão os gráficos $\log C(T, \epsilon, m)$ vs. $\log \epsilon$ para os processos estocásticos ARCH(2) e normal iid. Os gráficos apresentam nas partes lineares declividades crescentes com m variando de 2 a 5 (quadro 2.1). Assim, o algoritmo da dimensão de correlação está indicando corretamente que os processos observados não são caóticos (de baixa dimensão; geradores de números aleatórios são, em geral, processos caóticos com grande dimensão, por exemplo, 1.2.7, capítulo 1).

Nos exemplos anteriores, a dimensão de imersão máxima utilizada é $m=5$. As séries com comprimento $T=1500$ para esses sistemas não suportam dimensões de imersão superiores a $m=5$, isto é, o número de pontos em (2.8) com $\epsilon=1/2^i$ é muito pequeno ou nulo para $m>5$ e para valores de i que tornem ϵ suficientemente pequeno (na definição de dimensão de correlação (2.3), $\epsilon \rightarrow 0$).

QUADRO 2.1- DIMENSÃO DE CORRELAÇÃO ν PARA OS PROCESSOS ESTOCÁSTICOS NORMAL IID E ARCH(2)

m	PROCESSO	
	NORMAL IID	ARCH(2)
2	1.91(0.05)	1.85(0.03)
3	2.98(0.03)	2.74(0.03)
4	3.69(0.07)	3.61(0.07)
5	4.81(0.12)	4.39(0.07)

O valor entre parênteses é o erro padrão da estimativa da dimensão de correlação através da regressão (2.6).

Foram estimadas, também, a integral de correlação para os sistemas dinâmicos de Hénon e Lorenz, e de Hénon e Lorenz com

ruído: Hénon6, Hénon22, Lorenz1.5 e Lorenz12. O comprimento das séries para todos esses sistemas é $T=1500$.

A figura 2.3.a mostra os gráficos $\log C(T, \epsilon, m)$ vs. $\log \epsilon$ para a série Lorenz com $m=2, \dots, 7$ (para valores de $m > 7$ os gráficos se superpõem ao gráfico correspondente a $m=7$), e a figura 2.2.b mostra o gráfico da dimensão de correlação para $m=2, \dots, 10$. Observa-se que, na figura 2.2.b, a dimensão de correlação $\nu(T, m)$ tende a tornar-se independente da dimensão de imersão m , convergindo para a dimensão de correlação $\nu^*(T) \simeq 2.2$.

A figura 2.3.b apresenta os gráficos de $\log C(T, \epsilon, m)$ vs. $\log \epsilon$ para a série Lorenz1.5. Observa-se na figura 2.3.b, que os gráficos mudam de inclinação no intervalo $[-4.6, -3.6]$. Para $\log \epsilon = -4.6$, tem-se $\epsilon = 0.010$ e, para $\log \epsilon = -3.6$, tem-se $\epsilon = 0.027$, o que indica nível de ruído entre 1.0% e 2.7%. Para valores de $\log \epsilon$ menores que -3.6 , os gráficos mostram o efeito do ruído no sistema. Para valores de $\log \epsilon$ maiores que -3.6 e menores que -1.6 , os gráficos mostram que o sistema tem características caóticas, visto que os gráficos $\log C(T, \epsilon, m)$ vs. $\log \epsilon$ nas dimensões mais elevadas ($m=5, 6, 7$) têm partes lineares aproximadamente paralelas, indicando que a dimensão de correlação tende a tornar-se independente da dimensão de imersão, com m crescente.

A figura 2.3.c apresenta os gráficos de $\log C(T, \epsilon, m)$ vs. $\log \epsilon$ para a série Lorenz12. Para esse nível de ruído, os gráficos são dominados pelo efeito do ruído, reduzindo drasticamente a parte linear dos mesmos e dificultando as estimativas da dimensão de correlação para esse sistema.

As figuras 2.4.b e 2.4.c mostram os gráficos $\log C(T, \epsilon, m)$ vs. $\log \epsilon$ para as séries Hénon6 e Hénon22. Para a série Hénon6, os gráficos indicam, de forma análoga ao modelo de Lorenz, um sistema determinístico com ruído, mas para Hénon22 os gráficos são dominados pelo ruído, escondendo assim o comportamento caótico desse sistema.

As estimativas da dimensão de correlação dadas pelo algoritmo de Grassberg-Procaccia com (2.5) e (2.6), dependem da escolha visual das partes lineares dos gráficos $\log C(T, \epsilon, m)$ vs. $\log \epsilon$. A discussão a seguir procura mostrar a subjetividade desse procedimento e sugerir uma escolha menos arbitrária para o intervalo de variação de ϵ a ser considerado.

As figuras 2.5.a, 2.5.b e 2.5.c mostram os gráficos da dimensão de correlação para as séries Lorenz, Lorenz1.5 e Lorenz12, respectivamente, estimadas com o algoritmo de Brock-Baeck (2.7) e para os valores de $\epsilon = 2^{-i}$, $i=6, \dots, 16$. Para a série Lorenz, a dimensão de correlação nas dimensões mais elevadas, é aproximadamente constante no intervalo dado por $i=9$ a 11 , o que corresponde ao $\log \epsilon$ variando no intervalo $[-3.3, -2.7]$. Para a série Lorenz1.5 entretanto, o intervalo de variação de ϵ mais "apropriado" seria dado por $i=8$ a 10 , o que corresponde a

$\log \epsilon$ variando no intervalo $[-3.0, -2.4]$. A figura 2.5.c mostra que não é possível definir a dimensão de correlação para o sistema Lorenz12, a partir da correspondente série com comprimento $T=1500$.

Essa discussão procura esclarecer, também, como esse algoritmo perde sua eficiência em séries com nível de ruído significativo. Para sistemas com nível de ruído pequeno (veja Lorenz1.5, figura 2.5.b), o algoritmo exige séries com grande comprimento, de forma a aumentar a parte linear dos gráficos $\log C(T, m, \epsilon)$ vs. $\log \epsilon$. Os algoritmos da dimensão de correlação descartam observações; para séries com grande comprimento é possível escolher ϵ muito pequeno, o que tende a aumentar as partes lineares dos gráficos.

2.1.6 Limitações

As conclusões obtidas com a dimensão de correlação devem levar em consideração as seguintes limitações:

1) A dimensão de correlação exige séries com grande comprimento para identificar corretamente processos caóticos. A influência do comprimento da série no cálculo da dimensão de correlação é analisada por Theiler (1986, 1990) e Broomhead et al. (1989). Eles provam que um conjunto limitado de dados estocásticos ou determinísticos introduzem comportamento anômalo¹ na integral de correlação, inibindo boas estimativas. Smith (1988, p.283) prova que o número de pontos T necessário para estimar a dimensão de correlação com precisão de 5% é da ordem de $T_{\min} \geq 42^m$, onde m é o menor inteiro maior que a capacidade d_c do conjunto. Para um atrator com $d_c \approx 1.7$ (Mackey-Glass, por exemplo), $T_{\min} \geq 42^2 = 1764$ pontos.

2) A dimensão de correlação só pode ser observada para pequenos valores de ϵ (na definição (2.3), $\epsilon \rightarrow 0$). Como consequência, muitos valores da série são descartados pelo algoritmo, implicando em perda de informações. Para séries do mundo real em geral, com pequeno comprimento, esse fato representa um sério problema (Sugihara and May, 1990, p.738).

3) Existem limitações intrínsecas ao algoritmo. Ruelle (1989, p.245) argumenta que o intervalo de variação de ϵ no algoritmo de Grassberger-Procaccia deve ser da forma $r' \leq \epsilon \leq 10r'$. A partir desse intervalo, mostra que o algoritmo é limitado superiormente por $2 \log_{10} T$. Assim, dimensões de correlação com valores próximos ou maiores que a cota superior não devem ser consideradas².

4) Processos estocásticos podem ser identificados como

FIGURA 2.2: Gráficos das curvas $\log C(T, \epsilon, m)$ vs. $\log \epsilon$, estimadas através de (2.8, seção 2.1.3), para o sistema dinâmico de Mackey-Glass (fig. 2.2.a), com $T=1500$, $\epsilon=1/2^i$ e $m=2, \dots, 5$. As curvas mais baixas correspondem às dimensões de imersão mais elevadas. Gráfico da dimensão de correlação $\nu(T, \epsilon)$, estimado através de (2.6 e 2.8, seção 2.1.3), com $T=1500$, $\epsilon=1/2^i$ e $m=2, \dots, 10$, para o modelo de Lorenz (fig. 2.2.b).

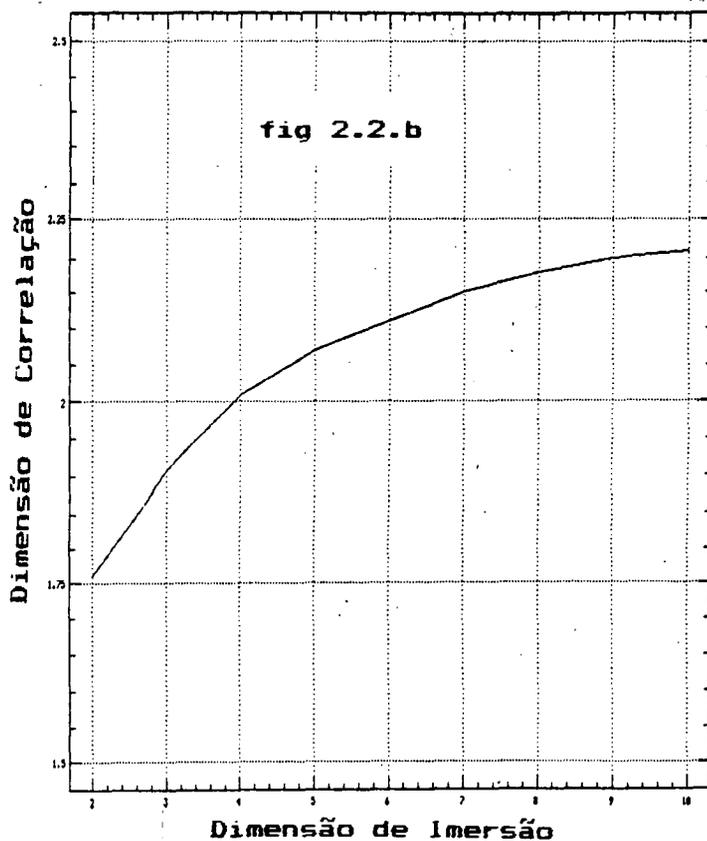
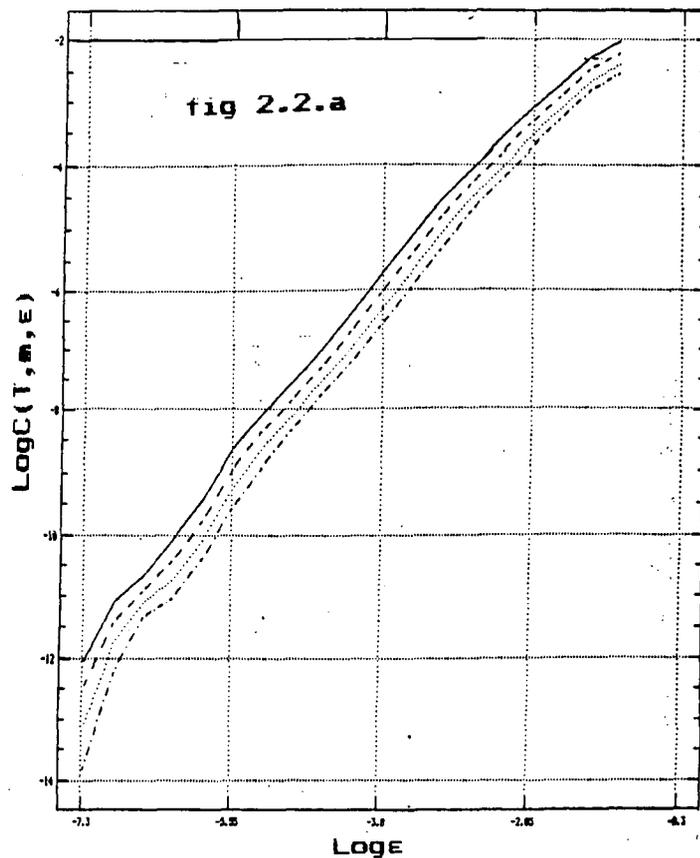


FIGURA 2.3. Gráficos das curvas $\log C(T, m, \epsilon)$ vs. $\log \epsilon$, estimados através de (2.8, seção 2.1.3), para o modelo de Lorenz (fig. 2.3.a), Lorenzi.5 (fig. 2.3.b) e Lorenzi2 (fig. 2.3.c) com $T=1500$, $\epsilon=1/2^i$ e $m=2, \dots, 7$. As curvas mais baixas correspondem às dimensões de imersão mais elevadas.

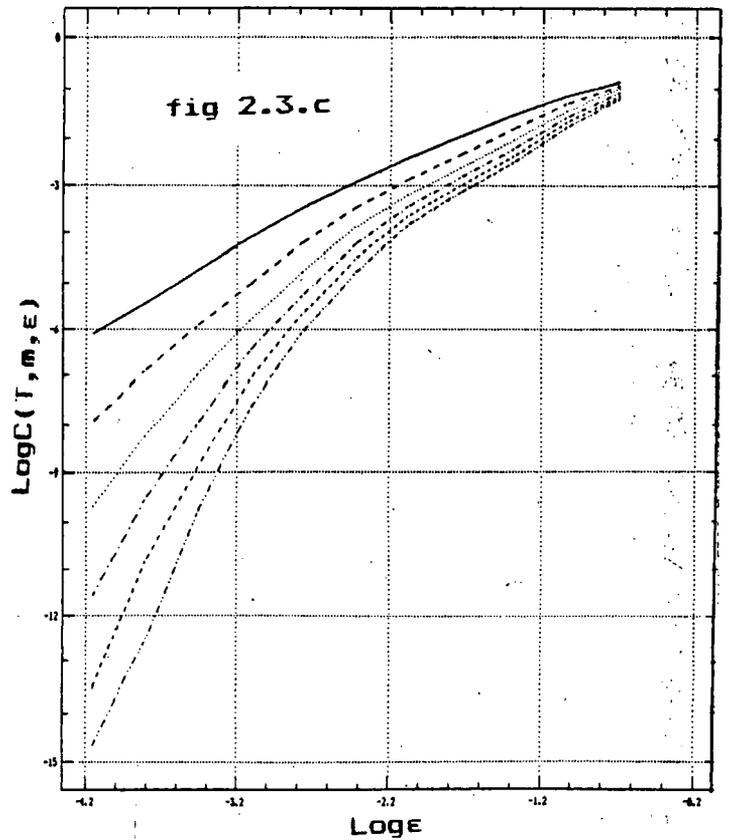
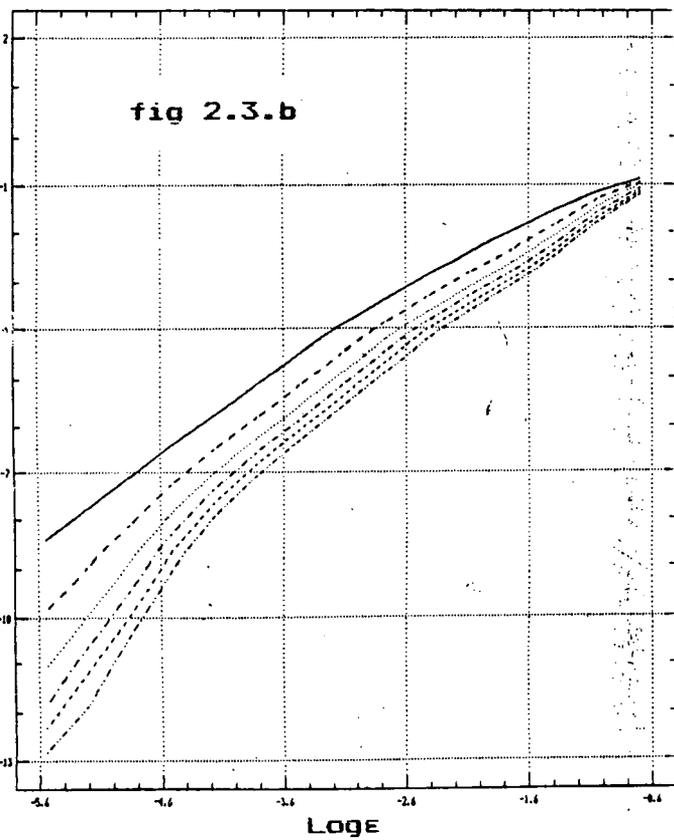
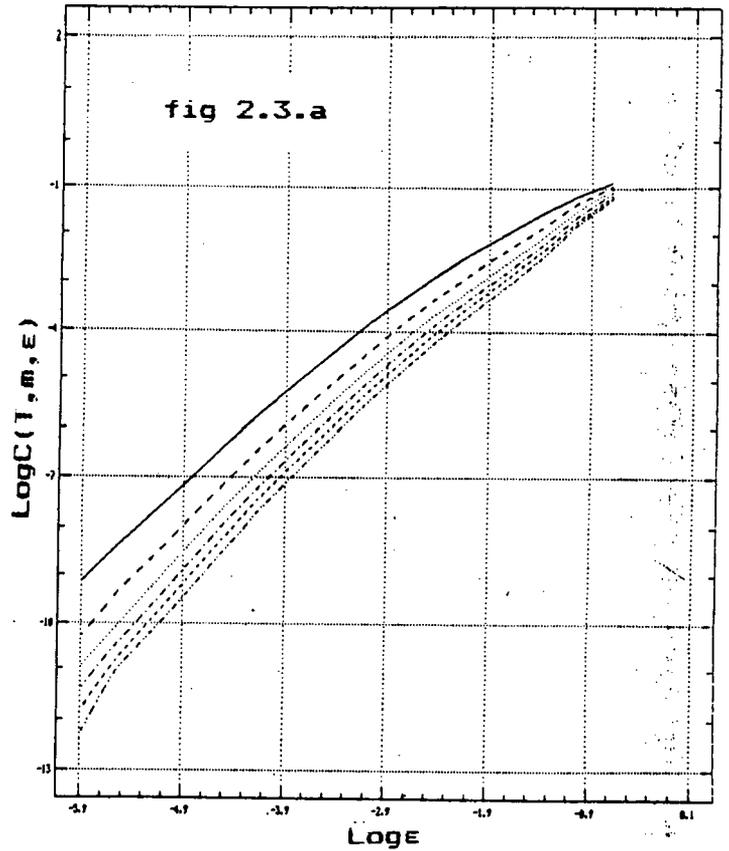


FIGURA 2.4. Gráficos das curvas $\log C(T, m, \epsilon)$ vs. $\log \epsilon$, estimadas através de (2.8, seção 2.1.3), para o sistema dinâmico da Hénon (fig. 2.4.a), Hénon6 (fig. 2.4.b) e Hénon22 (fig. 2.4.c), com $T=1500$, $\epsilon=1/2^i$ e $m=2, \dots, 7$. As curvas mais baixas correspondem às dimensões de imersão mais elevadas.

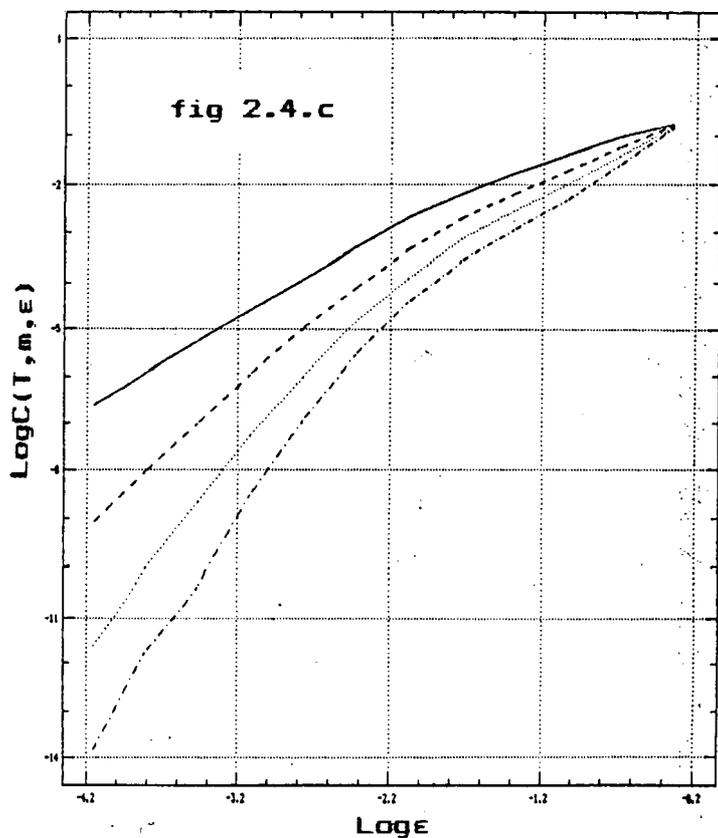
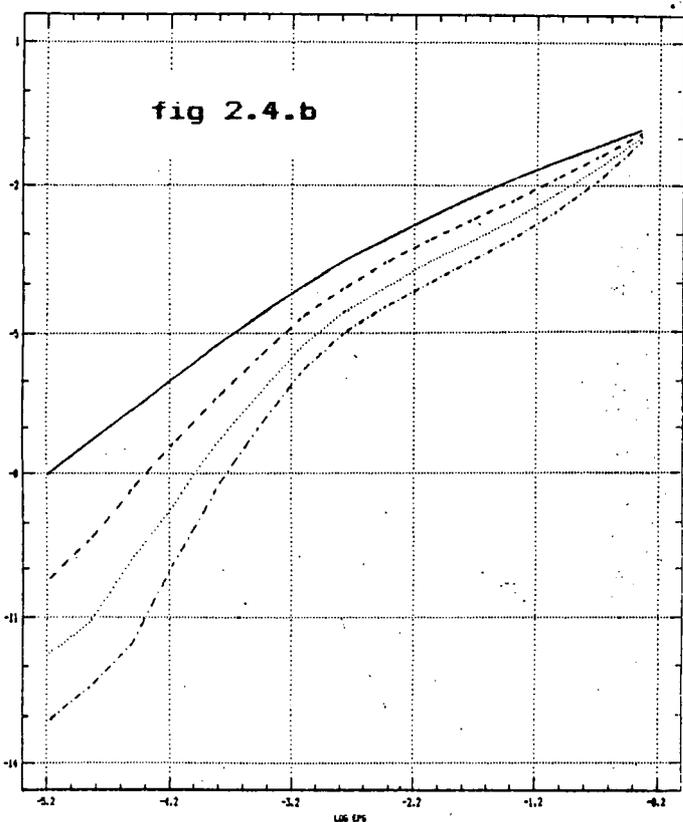
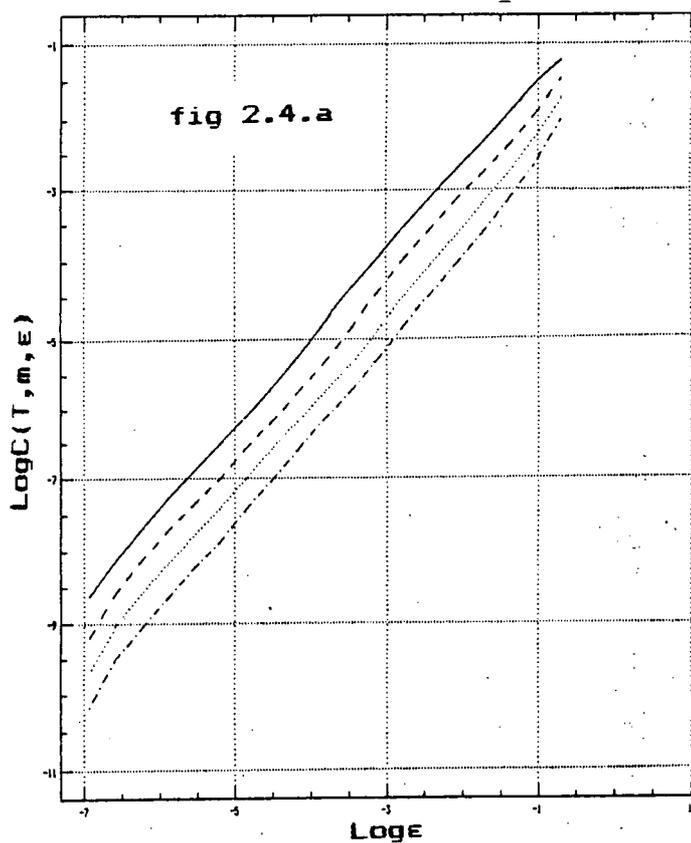


FIGURA 2.5. Gráficos da dimensão de correlação $\nu(T, \epsilon, \epsilon)$, estimadas através de (2.7 e 2.8, seção 2.1.3), para o modelo de Lorenz (fig. 2.5.a), Lorenzi1.5 (fig. 2.5.b) e Lorenzi2 (fig. 2.5.c), com $T=1500$, $\epsilon=1/2^i$, $i=6, \dots, 16$, $n=2, \dots, 7$.

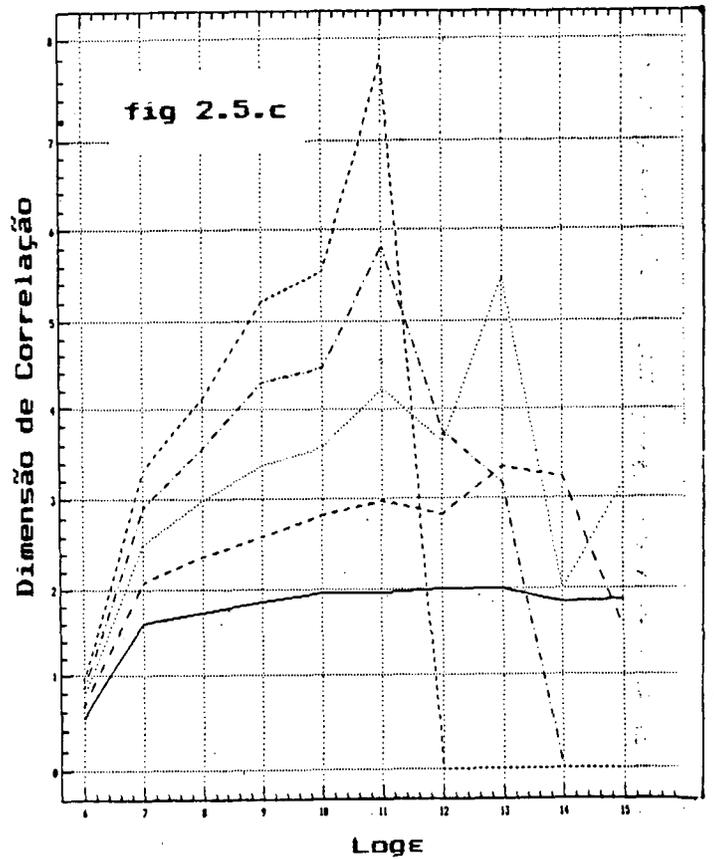
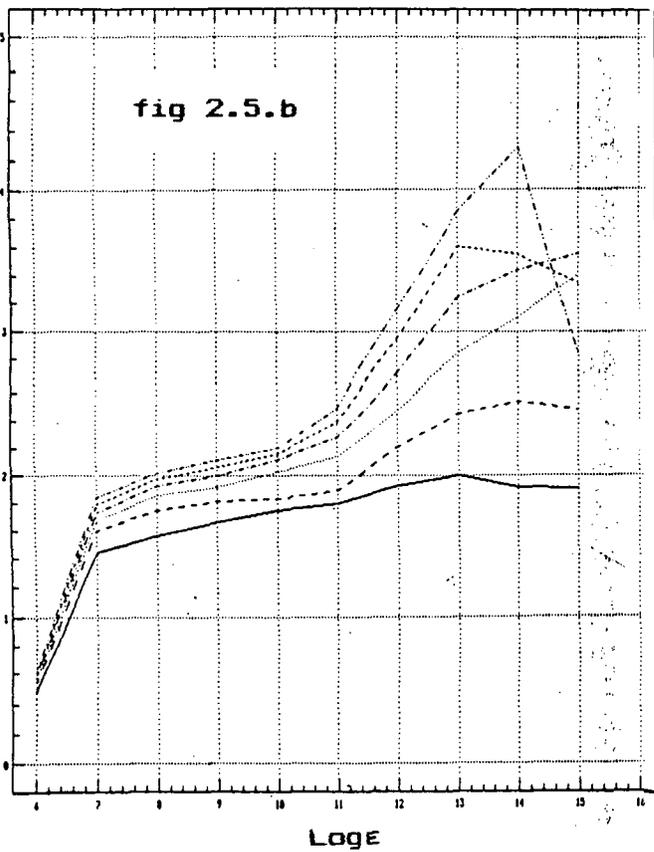
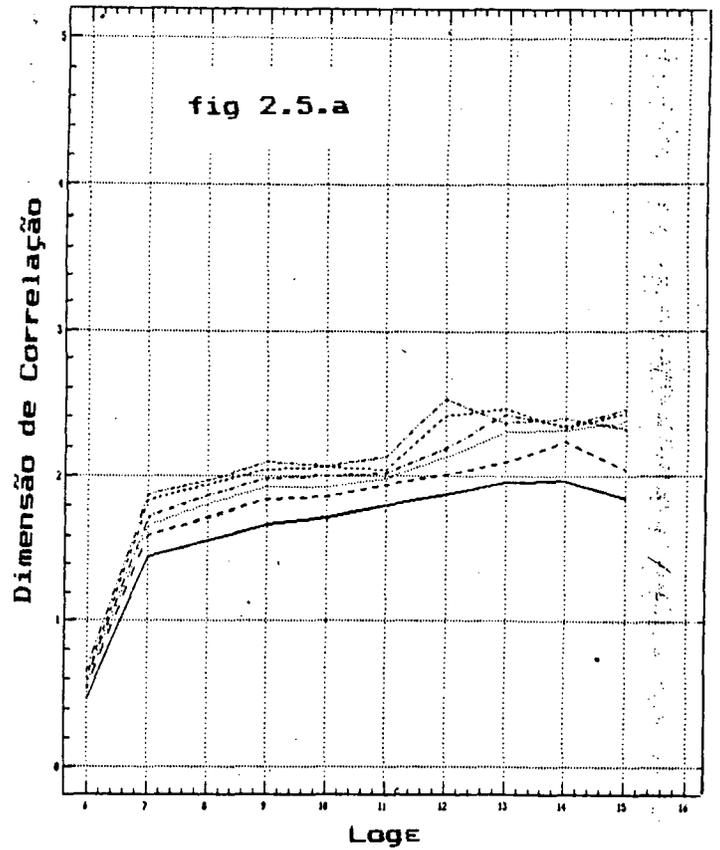
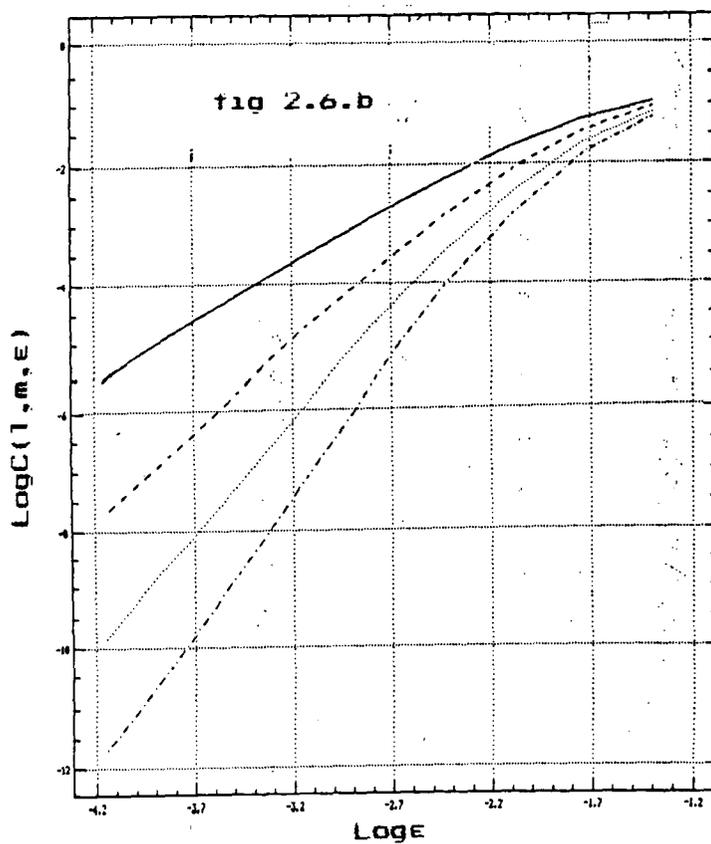
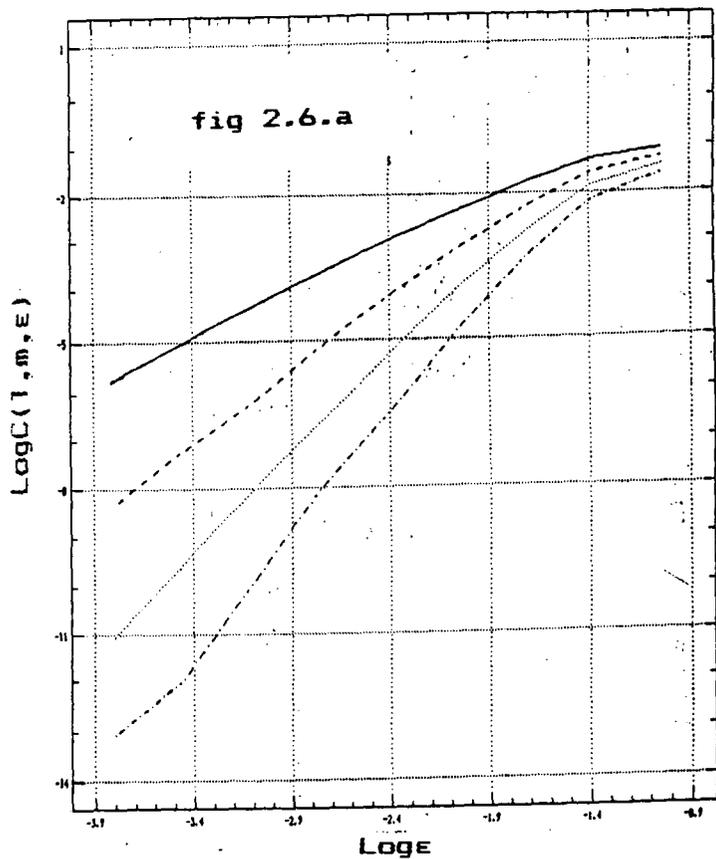


FIGURA 2.6. Gráficos das curvas $\log C(T, m, \varepsilon)$ vs. $\log \varepsilon$, estimadas através de (2.8, seção 2.1.3), para o processo normal iid. (fig. 2.6.a) e para o processo estocástico ARCH(2) (fig. 2.6.b), com $T=1500$, $\varepsilon=1/2^i$ a $m=2, \dots, 5$. As curvas mais baixas correspondem às dimensões da imersão mais elevadas.



caóticos. Brock and Sayers(1988,p.77) constroem um processo estocástico linear com raiz unitária que é identificado, através dos algoritmos da dimensão de correlação, como processo caótico.

As considerações acima permitem concluir que, para analisar séries do mundo real, são necessários algoritmos menos exigentes em número de dados e ao nível de ruído, nos quais a escolha de ϵ não seja importante para diagnosticar caos. Testes para distinguir sistemas caóticos de processos estocásticos que não dependam da escolha de ϵ e, portanto, menos exigentes quanto ao comprimento da série, são apresentados no capítulo 4.

2.2 EXPOENTES DE LYAPUNOV

Expoentes de Lyapunov generalizam, para sistemas não-lineares, os autovalores de sistemas lineares. De forma análoga aos autovalores, os expoentes de Lyapunov são empregados para analisar a estabilidade de sistemas dinâmicos. Sistemas não-lineares com expoente de Lyapunov positivo apresentam sensibilidade às condições iniciais, o que é uma marca característica dos processos caóticos. Este conceito é ampliado nesta seção para sistemas estocásticos não-lineares.

2.2.1 Definições

McCaffrey et al.(1992,p.653) definem expoentes de Lyapunov, para sistemas lineares, como segue:

¹ A dimensão de correlação é calculada com viés para cima para processos determinísticos caóticos e com viés para baixo para processos estocásticos(Ramsey and Yuan,1989,p.290).

² Scheinkman and LeBaron(1989,p.319), a partir de série financeira com $T=1227$, obtém dimensão de correlação próxima a 6(o limite dado pelo algoritmo é $2\log_{10} 1227=6.17$). A série permutada, explicação a seguir, apresenta dimensão de aproximadamente 11. Esses valores mostram, segundo Ruelle, que os limites do algoritmo não estão sendo respeitados.

Scheinkman and Lebaron fazem permutações aleatórias de séries para investigar processos caóticos. Argumentam que, para uma série gerada por processo aleatório e para a permutação dessa série, as estimativas da dimensão de correlação são muito próximas; enquanto que as estimativas da dimensão de correlação para uma série gerada por sistema determinístico caótico e para a permutação dessa série, são muito diferentes(procedimento análogo a esse é empregado por Ramsey et al.,1990,p.998).

Considere-se o modelo linear auto-regressivo

$$x_t = \sum_{k=1}^d a_k x_{t-k} + \varepsilon_t, \quad (2.9)$$

onde ε_t é iid. com $E(\varepsilon_t)=0$ e $\text{VAR}(\varepsilon_t)=\sigma^2$. O modelo (2.9) pode ser escrito em termos de variáveis de estado na forma

$$X_t = AX_{t-1} + E_t, \quad (2.10)$$

onde A é uma matriz $d \times d$ e $X'_t, E'_t \in \mathbb{R}^d$. Os autovalores $\gamma_i, i=1, \dots, d$, do sistema determinístico linear

$$X_t = AX_{t-1} \quad (2.11)$$

associado ao sistema linear auto-regressivo (2.10), são dados por

$$\det(A - I\gamma) = 0.$$

Os expoentes de Lyapunov correspondentes aos autovalores γ_i são os números

$$\lambda_i = \log(|\gamma_i|). \quad (2.12)$$

O conjunto dos expoentes de Lyapunov (2.12) ordenados em ordem decrescente

$$\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_d,$$

é chamado de espectro de Lyapunov. O maior expoente de Lyapunov λ_1 tem papel de destaque na análise de sistemas dinâmicos e é chamado de expoente de Lyapunov dominante.

É possível mostrar que para toda condição inicial X_0

$$\lim_{t \rightarrow \infty} (1/t) \log \|A^t X_0\| \rightarrow \max\{\lambda_i, i=1, \dots, d\}, \quad (2.13)$$

isto é, o limite em (2.13) é o maior expoente de Lyapunov do sistema linear (2.11) (McCaffrey et al., 1992, p.683). Se $\lambda_1 > 0$, todas as raízes γ_i são exteriores ao círculo unitário e as trajetórias de (2.11) divergem exponencialmente. Se $\lambda_1 < 0$, todas as raízes γ_i são interiores ao círculo unitário e as trajetórias convergem exponencialmente para zero. McCaffrey et al. (1992,

p.684) mostram que esse comportamento é transportado pelo sistema linear (2.11) para o sistema linear estocástico (2.10). Se $\lambda_1 < 0$, as trajetórias são oscilações estacionárias com variância limitada, enquanto que para $\lambda_1 > 0$, a variância diverge e não há convergência para distribuição estacionária.

2.2.2 Sistemas dinâmicos não-lineares

As definições em 2.2.1 são generalizadas para sistemas dinâmicos não-lineares por Nychka et al. (1992, p.405):

Considere-se o modelo auto-regressivo não-linear

$$x_t = f(x_{t-1}, x_{t-2}, \dots, x_{t-p}) + \varepsilon_t, \quad (2.14)$$

onde ε_t é iid. com $E(\varepsilon_t) = 0$, $\text{VAR}(\varepsilon_t) = \sigma^2$ e $f: \mathbb{R}^m \rightarrow \mathbb{R}$ não-linear. O modelo auto-regressivo (2.14) pode ser escrito em termos de variáveis de estado na forma³

$$X_t = F(X_{t-1}) + E_t, \quad (2.15)$$

onde $X'_t, E'_t \in \mathbb{R}^m$ e $F: \mathbb{R}^d \rightarrow \mathbb{R}^d$ uma aplicação não-linear.

O expoente de Lyapunov dominante λ_1 do sistema não-linear auto-regressivo (2.15) é definido, de forma análoga ao sistema linear (2.13), por

$$\lambda_1 = \lim_{t \rightarrow \infty} (1/t) \log \| J_{t-1} * J_{t-2} * \dots * J_0 \|, \quad (2.16)$$

onde J_t é a matriz das derivadas parciais da aplicação F em (2.15) calculada em X_t , isto é, $J_t = DF(X_t)$.

Para os sistemas dinâmicos não-lineares, o espectro de Lyapunov depende, em geral, da condição inicial X_0 . O teorema de Oseledec (MacCaffrey et al., 1992, p.684) estabelece as condições que permitem definir o espectro de Lyapunov para sistemas dinâmicos não-lineares. Em particular, afirma que para os

³ Em Priestley (1980) vários tipos de modelos não-lineares, representados em termos de variáveis de estado, são apresentados. Nessa referência são discutidos também algoritmos para a identificação e para a previsão com esses modelos.

sistemas ergódicos⁴ o maior expoente de Lyapunov é independente das condições iniciais X_0 (MacCaffrey et al., 1992, p.684).

Para os sistemas determinísticos caóticos, trajetórias com condições iniciais arbitrariamente próximas divergem exponencialmente. Essa propriedade, conhecida como sensibilidade às condições iniciais, é a marca característica da dinâmica caótica e caracteriza a imprevisibilidade dos sistemas caóticos no longo prazo. A taxa média de divergência das trajetórias é dada pelo expoente de Lyapunov dominante λ_1 . Se $\lambda_1 > 0$, trajetórias com condições iniciais arbitrariamente próximas divergem, enquanto que para $\lambda_1 < 0$, essas trajetórias convergem.

Expoente de Lyapunov positivo em sistema estocástico indica sensibilidade às condições iniciais, mostrando que não-linearidades contribuem para a imprevisibilidade das trajetórias do sistema dinâmico (Nychka et al., 1992, p.405). Yao and Tong (1994, p.704) procuram caracterizar sistemas estocásticos com sensibilidade às condições iniciais:

Operacionalmente, os sistemas dinâmicos (2.15) têm sensibilidade às condições iniciais (são sistemas estocásticos caóticos) se:

1) para algum $N \geq 1$ (N é o número de iterações), a distribuição condicional de X_1, \dots, X_N , dada a condição inicial $X_0 = x$, depende sensivelmente dessa condição para alguns valores de x ;

2) para todo $t \geq 1$, o ruído E_t é amplificado muito rapidamente (?) a partir de alguma condição inicial X_0 , através da evolução da dinâmica do sistema no tempo.

2.2.3 Algoritmos para estimar o expoente de Lyapunov dominante

Dada a série temporal $\{x_t, t=1, \dots, T\}$ e a dimensão de imersão m , a dinâmica do sistema pode ser reconstruída pelo método das defasagens

$$X_t = (x_t, x_{t-\tau}, \dots, x_{t-(m-1)\tau}).$$

⁴ Sistemas para os quais os invariantes geométricos e dinâmicos podem ser estimados a partir de uma trajetória do sistema no tempo, obtida a partir de uma condição inicial típica qualquer, são sistemas ergódicos.

Além disso, o teorema de Takens afirma que existe f não-linear diferenciável

$$f: \mathbb{R}^m \rightarrow \mathbb{R}^m,$$

tal que

$$x_{t+1} = f(x_t, x_{t-\tau}, \dots, x_{t-(m-1)\tau}). \quad (2.17)$$

A partir da reconstrução da dinâmica do sistema, métodos diretos como os de Wolf et al.(1985), comparam as trajetórias e avaliam como elas divergem, de modo a obter uma taxa média de divergência, que é uma estimativa do expoente de Lyapunov dominante. Métodos Jacobianos, tais como os de Eckmann and Ruelle(1985), Eckmann et al.(1986) e Sano and Sawada(1985), estimam a matriz Jacobiana $J_t(x_t)$ em (2.15), através de aproximações locais da função f em (2.17). O expoente de Lyapunov dominante é, então, estimado por(Nychka,1992,p.407):

$$\lambda_1 = (1/2T) \log(|\gamma_1(T)|), \quad (2.18)$$

onde $\gamma_1(T)$ é o maior autovalor de $J'(T) * J(T)$, com

$$J(T) = J(x_{t-1}) * J(x_{t-2}) * \dots * J(x_0).$$

A relação em (2.18) é uma versão numericamente mais estável de (2.16), para T finito.

2.2.4 Exemplos numéricos⁵

Com os algoritmos de Wolf et al.(1985)(WSSV), Jacobianos com aproximações locais(LLOCAL) e aproximações globais(GLOBAL)^{6,7} foram estimados os expoentes de Lyapunov dominantes para os sistemas determinísticos caóticos da Logística, Hénon, Mackey-Glass, Ikeda e Lorenz. Os comprimentos das séries são: $T=1000$ para o algoritmo WSSV, $T=200$ para o algoritmo GLOBAL e $T=1500$ para o algoritmo LOCAL. Para o algoritmo GLOBAL, foram

⁵ Os programas computacionais e arquivos(séries) de dados utilizados nos diversos capítulos desta tese podem ser obtidos com o autor no endereço:

Universidade Federal do Paraná-Departamento de Matemática
 Telefone (041) 366-2323, Ramal 3126
 Caixa Postal 19081
 CEP 81531-990, Curitiba - Paraná

construídas 500 aproximações locais das matrizes Jacobianas de cada sistema dinâmico. A dimensão de imersão (número de defasagens na equação (2.17)), para a reconstrução da dinâmica do sistema, foi obtida com o algoritmo de modelagem apresentado no capítulo 4.

No quadro 2.2, estão os resultados obtidos. Os valores positivos, dos expoentes de Lyapunov dominantes estimados através dos três algoritmos confirmam, para esses sistemas, a sensibilidade às condições iniciais, característica essa dos sistemas determinísticos caóticos.

As estimativas obtidas com o algoritmo GLOBAL mostram a possibilidade de identificar sistemas determinísticos caóticos com séries com pequeno comprimento (Dechert and Gencay, 1992, p. S59), contornando a exigência de séries com grande comprimento para identificar caos através das estimativas da dimensão de correlação ou do expoente de Lyapunov dominante, com os métodos diretos e Jacobianos com aproximações locais.

QUADRO 2.2 - EXPOENTE DE LYAPUNOV DOMINANTE. SISTEMAS DETERMINÍSTICOS CAÓTICOS

SISTEMA	ALGORITMO		
	WSSV	LOCAL	GLOBAL
Logística	0.6800	0.6919	0.6932
Hénon	0.4273	0.4102	0.4067
M.-Glass	0.1403	0.1307	0.1209
Ikeda	0.3070	0.3658	0.3369
Lorenz	0.1696	0.2282	0.2292

Os comprimentos das séries são: T=1000 para WSSV, T=1500 para LOCAL e T=200 para GLOBAL.

Foram, também, estimados os expoentes de Lyapunov dominantes para os sistemas com ruído Hénon⁶, Hénon²², Lorenz^{1.5}, Lorenz¹², e para os processos estocásticos ARCH(2) e das Manchas Solares (nota⁵, capítulo 3). Os comprimentos das séries são os mesmos dos correspondentes sistemas determinísticos, exceto para

⁶ No capítulo 3, as regressões utilizadas neste texto para obter aproximações locais e globais da função f são apresentadas.

⁷ Para séries com pequeno comprimento, a matriz Jacobiana pode ser estimada através de aproximações globais de f , como foi feito por Dechert and Gencay (1992, p. S44).

a série das Manchas Solares que é de comprimento $T=280$. No quadro 2.3, estão as estimativas obtidas com os três algoritmos.

As estimativas para os sistemas não-lineares com ruído(quadro 2.3) mostram que esses sistemas apresentam sensibilidade às condições iniciais, no sentido introduzido por Nychka et al.(1992). Para o sistema de Hénon, o nível de ruído de 6% tem pequeno efeito no comportamento caótico do sistema. O expoente de Lyapunov estimado(algoritmo GLOBAL) para o modelo

QUADRO 2.3-EXPOENTE DE LYAPUNOV DOMINANTE. SISTEMAS NÃO-LINEARES COM RUÍDO

SISTEMA	ALGORITMO		
	WSSV	LOCAL	GLOBAL
Hénon6	0.3729	0.3919	0.4173
Hénon22	0.2767	0.2173	0.2665
Lorenz1.5	0.1223	0.1193	0.1329
Lorenz12	0.0684	0.0520	0.0574
M. Solares	0.0595	-0.1006	0.0272
ARCH(2)	0.0133	0.0170	-0.4061

Os comprimentos das séries são: $T=1000$ para WSSV, $T=1500$ para LOCAL e $T=200$ para GLOBAL. Para a série M. Solares, $T=280$.

determinístico puro é 0.4067, enquanto que para o modelo com nível de ruído de 6%, é estimado em 0.4173.

Para os sistemas caóticos com ruído, o expoente de Lyapunov é, em geral, menor que o do correspondente sistema determinístico (quadros 2.2 e 2.3). O efeito do ruído é causar o "sombreamento" do comportamento caótico, escondendo o caos(Ben-Mizrachi et al.,1984,p.976).

Para o sistemas estocásticos ARCH(2) e Manchas Solares, os expoentes de Lyapunov estimados pelo algoritmo WSSV são, respectivamente, $\lambda_1=0.0595$ e $\lambda_1=0.0133$.

Um dos problemas de diagnosticar caos, tendo como critério apenas o expoente de Lyapunov dominante, fica claro nesses dois últimos exemplos. Os expoentes estimados são significativos? Não há resposta afirmativa para essa pergunta. As distribuições estatísticas para o expoente de Lyapunov e, também, para a dimensão de correlação, não são conhecidas(Scheikman,1990,p.47; Brock and Baek,1991,p.697).

No capítulo 4, é discutida a adequação dos algoritmos para as estimativas do expoente de Lyapunov dominante e, também, da dimensão de correlação para dada série temporal.

2.2.5 Limitações

As conclusões obtidas através do expoente de Lyapunov devem levar em consideração as seguintes limitações:

1) Os métodos Jacobianos com aproximações locais exigem séries com grande comprimento e com baixo nível de ruído (McCaffrey et al., 1990, p.655). Abarbanel et al. (1990, p.1790) sugerem, como regra geral, que o número de dados necessários para esse tipo de algoritmo é da ordem de 20^m , onde $m > 1$ é a dimensão de imersão. Por exemplo, para o sistema de Hénon, na dimensão de imersão $m=3$, tem-se que o número de dados necessários é da ordem de $20^3=8000$.

2) A função f , dada pelo teorema de Takens, depende da dimensão de imersão m utilizada para a reconstrução do sistema dinâmico. Ela determina a dimensão do modelo, isto é, o número de defasagens em (2.17). Em geral, a dimensão do modelo não é conhecida⁸, obrigando a estimar f com m crescente até que o expoente de Lyapunov dominante torne-se estável. Se a técnica de regressão para estimar f é sensível a defasagens incorretas, esse procedimento não pode ser aplicado (Nychka et al., 1992, p.412)⁹.

3) Cheng and Tong (1992, p.439) apontam outro problema no procedimento de variar a dimensão de imersão, comum em muitos algoritmos para diagnosticar caos em séries temporais:

O procedimento de variar a dimensão de imersão implica em perda de informações dadas por séries temporais de comprimento limitado. O procedimento correto é determinar em primeiro lugar a dimensão de imersão e depois a estatística de interesse.

Os algoritmos de Wolf et al. (1985) (WSSV) e Eckmann et al. (1986), para estimar expoentes de Lyapunov, e o de Grassberger-Procaccia (1983a), para estimar a dimensão de correlação, são exemplos de algoritmos que empregam dimensão de imersão crescente para diagnosticar caos.

⁸ No capítulo 4, o problema de determinar a dimensão de imersão correta para dada série temporal é discutido.

⁹ Nychka et al. (1992, p.407) e McCaffrey et al. (1992, p.687) analisam diversos tipos de regressão não-paramétrica para estimar f , concluindo que as regressões não-paramétricas dadas pelas redes neurais são as mais robustas em relação a escolha incorreta da dimensão do modelo.

4) Os métodos diretos, implementados através da comparação de trajetórias similares, são muito sensíveis ao nível de ruído na série temporal. Como colocado por Nychka et al. (1992, p.407):

As trajetórias que estão sendo comparadas podem não ter a mesma distribuição de choques aleatórios. Assim, a divergência entre elas pode ser devida, simplesmente, aos choques. Além disso, os métodos diretos exigem séries com grande comprimento.

O número de dados recomendados por Wolf et al. (1985, p.305) para o algoritmo WSSV é da ordem de 30^m , onde $m > 1$ é a dimensão de imersão. Para o sistema dinâmico de Hénon, tem-se na dimensão de imersão $m=3$, que o número de pontos recomendado é da ordem de $30^3=27000$.

As regras práticas para determinar o número de dados necessários aos algoritmos Jacobianos (20^m) e diretos (30^m), presupõem que a distribuição dos pontos no atrator é uniforme e que não existe ruído. Atratores com distribuição de pontos não uniforme e com ruído exigem, em geral, maior quantidade de dados do que as sugeridas por essas regras.

As colocações acima permitem concluir que diagnosticar caos, tendo como critério o expoente de Lyapunov dominante, estimado através de métodos diretos ou Jacobianos em séries com pequeno comprimento e com perturbações estocásticas não negligenciáveis, como é o caso da séries do mundo real, podem dar resultados incorretos ou ambíguos.

No capítulo 4, outras técnicas, empregando modelos de previsão, são discutidas para diagnosticar caos em séries temporais.

CAPÍTULO 3

PREVISÃO DE SÉRIES TEMPORAIS NÃO-LINEARES

A análise de séries temporais não-lineares, apresentada no capítulo 2, não aborda o problema fundamental em séries temporais com relação à previsão. Esse problema pode ser tratado adequadamente com modelos obtidos por regressão não-paramétrica. Neste capítulo, são apresentadas técnicas de regressão não-paramétrica para sistemas auto-regressivos não-lineares, dos quais os sistemas determinísticos caóticos são um caso particular.

3.1 MODELOS DE PREVISÃO

A análise de séries temporais não-lineares, através da dimensão de correlação e do expoente de Lyapunov, não aborda o problema fundamental em séries temporais com relação à previsão. As estimativas desses invariantes geométricos e dinâmicos, juntamente com os testes estatísticos não-lineares, são de utilidade limitada para a construção de modelos de previsão. Conclui-se que a série temporal é proveniente de um modelo determinístico caótico ou de um modelo estocástico não-linear, de modo que, em princípio, é possível obter para ela um modelo de previsão. O maior expoente de Lyapunov indica como os erros de previsão crescem no tempo. A dimensão de correlação indica qual é a complexidade do modelo, isto é, o número mínimo de variáveis necessárias para modelar o sistema. As estatísticas não-lineares apontam que o modelo apropriado é não-linear. Casdagli(1989,p. 335) argumenta que:

Esse tipo de análise é de uso restrito, pois não fornece informações que permitam construir um modelo de previsão para a série temporal gerada pelo sistema dinâmico e apresenta técnicas numéricas para construir modelos não-lineares de previsão diretamente da série temporal. Argumenta, também, que é preferível um modelo aproximado de previsão, monitorado pelo erro cometido, do que estimativas precisas da dimensão de correlação ou do expoente de Lyapunov.

3.1.1 Tipos de Modelos

Modelos locais

Modelos locais de previsão¹ para séries determinísticas caóticas surgem na literatura com o artigo de Farmer and Sidorowich(1987) e Casdagli(1989). Partindo do teorema de Takens, que afirma existir f não-linear tal que para todos inteiros p , τ e para a dimensão de imersão m ,

$$x_{t+p} = f^p(x_t), \quad (3.1)$$

onde

$$X_t = (x_t, x_{t-\tau}, \dots, x_{t-(m-1)\tau}), \quad (3.2)$$

são os vetores que reconstróem a dinâmica do sistema a partir da série temporal observada $\{x_t: t=1, \dots, T\}$, obtêm-se aproximações locais lineares de f^p , que são utilizadas para fazer a previsão \hat{x}_{t+p} de x_t , p passos à frente.

Um modelo de previsão local, baseado em simplex m -dimensional, é construído por Sugihara et al.(1990), Sugihara and May(1990,p.736) para distinguir caos de ruído em séries temporais da Biologia.

Modelos globais

Nychka et al.(1992,p.401) e Dechert and Gencay(1992,p.543) consideram sistemas estocásticos não-lineares

$$x_{t+1} = f(X_t) + \epsilon_{t+1}, \quad (3.3)$$

e aproximam f globalmente através de redes neurais. A previsão \hat{x}_{t+p} de x_t , p passos à frente, com f aproximada globalmente é dada por

$$\hat{x}_{t+p} = f^p(X_t) = (f^{p-1} \circ f)(X_t) \quad (3.4)$$

onde "o" indica a "composição de funções".

¹ Enfoque diferente do aqui apresentado é desenvolvido em Arbabanel et al.(1989,1990). Empregando o expoente de Lyapunov e a densidade invariante(Arbabanel et al,1980,p.1784), como restrições, obtém-se a função f , a qual modela a dinâmica do sistema e reproduz esses invariantes.

Os modelos de previsão locais de Farmer and Sidorowich e Casdagli podem ser implementados com regressões não-paramétricas locais disponíveis na literatura econométrica (seção, 3.2), enquanto que os modelos de previsão global podem explorar a habilidade que as redes neurais têm de aproximar funções e suas derivadas com ótima precisão (seção 3.5).

Os modelos locais e globais de previsão têm excelente desempenho para séries caóticas com comprimento moderado (quando comparada com as séries exigidas para estimar o expoente de Lyapunov dominante e a dimensão de correlação pelos algoritmos WSSV e Grassberger-Procaccia) e com baixo nível de ruído (Farmer and Sidorowich, 1987; Casdagli, 1989).

Neste trabalho, pretende-se investigar a desempenho desses modelos em sistemas estocásticos não-lineares

$$x_{t+1} = f(x_t) + \epsilon_{t+1}, \quad (3.5)$$

x_t dado em (3.2), dos quais os sistemas determinísticos caóticos são um caso particular ($\epsilon_t=0$), para séries com pequeno comprimento (nota 9, capítulo 1) e nível de ruído significativo (1.2.8, capítulo 1).

A validação indireta dos modelos de previsão através dos erros de previsão e da reprodução de invariantes geométricos e dinâmicos conhecidos (dimensão de correlação e expoente de Lyapunov por exemplo), para séries com grande comprimento e baixo nível de ruído (como é feito por Farmer and Sidorowich e Casdagli), será substituída por testes estatísticos apropriados nos resíduos dos modelos obtidos para as séries com pequeno comprimento e com nível de ruído significativo.

3.2 REGRESSÃO NÃO-PARAMÉTRICA LOCAL

Nesta seção, são apresentadas as técnicas de regressão não-paramétricas empregadas neste trabalho para aproximar a função f em (3.1) dada pelo teorema de Takens. Essas aproximações permitem implementar os algoritmos Jacobianos para estimar o expoente de Lyapunov dominante e construir modelos de previsão para séries temporais não-lineares. A modelagem de séries temporais não-lineares (capítulo 4) será feita, também, através destas regressões.

3.2.1 Nearest-neighbor

Considere o modelo de regressão

$$y_t = f(X_t) + \varepsilon_t, \quad (3.6)$$

onde $X_t \in \mathbb{R}^p$, $t=1, \dots, T$, $f: \mathbb{R}^p \rightarrow \mathbb{R}$ é uma função não-linear e $\{\varepsilon_t\}$ é iid. com $E(\varepsilon_t | X_t) = 0$.

A função f pode ser estimada através da técnica de regressão não-paramétrica Nearest-neighbor²(NN)(Robinson, 1983; Yakowitz 1987).

Segundo Mizrach(1992,p.S153), a estimativa de f através da técnica NN em um ponto arbitrário $X^* \in \mathbb{R}^p$ é dada por:

$$\hat{f}(X^*) = \sum_{t=1}^T w_t I_1(X_t - X^*) y_t, \quad (3.7)$$

onde I_1 é função indicadora

$$I_1 = \begin{cases} 1, & \text{se } \|X_t - X^*\| < l \\ 0, & \text{se } \|X_t - X^*\| \geq l, \end{cases} \quad (3.8)$$

l é uma constante(dita constante de alisamento) e w_1, \dots, w_T é uma sequência de pesos.

A escolha de l em (3.8) determina o número k de pontos usados em cada vizinhança de X^* , para estimar o valor da regressão em (3.7).

Para que as estimativas de f sejam consistentes, é necessário e suficiente que a cardinalidade k do conjunto de pontos, dados pela função indicadora, tenha as propriedades

$$\lim_{T \rightarrow \infty} k = \infty \quad \text{e} \quad \lim_{T \rightarrow \infty} k/T = 0, \quad (3.9)$$

isto é, que o número de pontos em cada vizinhança seja crescente com T , mas com taxa menor do que a dada pelo crescimento da amostra(Stone, 1977, p.600).

Com as condições acima, os pesos podem ser escolhidos através de vários esquemas: uniforme, quadrático ou triangular, etc³. Pesos uniformes são dados por

$$w_t = \begin{cases} 1/k, & \text{se } X_t \text{ é ponto na vizinhança de } X^*, \\ 0, & \text{caso contrário.} \end{cases} \quad (3.10)$$

² Para a formulação geral deste problema, ver Priestley and Chao(1972) e Granger and Newbold(1976).

3.2.2 Locally-weighted regression

A regressão Locally-weighted Regression (LWR)⁴ (Cleveland and Delvin, 1988) é uma generalização da regressão NN. De forma semelhante a técnica NN, LWR utiliza pontos na vizinhança de X^* para estimar a regressão. Mas a média em (3.7) é substituída pela regressão linear de y_t em X_t^* , X_t^* os pontos na vizinhança de X^* dados pela função indicadora (3.8).

A estimativa LWR do valor da regressão $f(X^*)$, no ponto $X^* \in \mathbb{R}^p$, é obtida como se segue (Diebold and Nason, 1990, p.319):

Para dada constante de alisamento l , seja k o número de pontos X_t^* na vizinhança de X^* dados pela função indicadora I_l , ordenados segundo suas distâncias ao ponto X^* , isto é, $X_1^*, X_2^*, \dots, X_k^*$, são tais que X_i^* é o i -ésimo ponto mais próximo de X^* . Seja

$$d(X^*, X_k^*) = \left[\sum_{j=1}^k (x_{kj}^* - x_j^*)^2 \right]^{1/2}, \quad (3.11)$$

a distância euclidiana de X^* ao ponto X_k^* . Construa-se a função peso

$$w_t(X_t, X^*, X_k^*) = h(d(X_t, X^*)/d(X^*, X_k^*)), \quad (3.12)$$

onde h é, por exemplo, a função tricúbica (Cleveland, 1979, p.831):

$$h(u) = \begin{cases} (1 - u^3)^3, & \text{se } u < 1, \\ 0, & \text{se } u \geq 1. \end{cases} \quad (3.13)$$

O valor da regressão linear no ponto X^* é, então, computado por:

$$\hat{y}^* = \hat{f}(X^*) = \hat{\alpha}_0 + X^* \hat{\beta}, \quad (3.14)$$

³ Cleveland (1979) e Diebold and Nason (1990) recomendam que os pesos sejam escolhidos através da função tricúbica (3.13). Outros esquemas para os pesos, incluindo pesos quadráticos e triangulares, são dados em Priestley and Chao (1972, p.390).

⁴ Tratamento rigoroso do estimador LWR pode ser encontrado em Cleveland et al. (1988). Métodos de regressão local, propriedades e algoritmos são discutidos nessa referência.

onde $\hat{\alpha}_0$ e $\hat{\beta}$ minimizam

$$\sum_{t=1}^T w_t (y_t - \alpha_0 - X_t' \beta)^2. \quad (3.15)$$

Esse problema de mínimos quadrados pode ser resolvido eficientemente pelo algoritmo QR (Decomposição de uma matriz real A na forma $A=Q \cdot R$, onde Q é ortogonal e R é triangular superior) (Press et al., 1986, p.369). O número de pontos k necessário para dar estabilidade numérica ao algoritmo QR, com m variáveis explicativas, é de no mínimo $m+1$. Experiências computacionais com séries determinísticas caóticas mostram que valores de $k \geq 2(m+1)$, onde m é a dimensão de imersão, torna o algoritmo numericamente mais estável, conduzindo a melhores resultados.

De forma análoga à regressão não-paramétrica NN, as estimativas locais de f obtidas com a regressão LWR são consistentes se valem as relações estabelecidas em (3.9).

O procedimento LWR descrito acima pode ser modificado. A distância euclidiana (3.11) pode ser substituída por outra norma no espaço R^p , os pesos (3.13) podem ser obtidos através de outros esquemas (Mizrach, 1992, p.5154) e a regressão linear local (3.14) pode ser substituída por regressão local de ordem maior, tal como a quadrática. O valor mínimo de k (número de pontos) depende da ordem da regressão local.

Considerando (3.2) e $y_t = x_{t+1}$ em (3.6), tem-se

$$x_{t+1} = f(X_t) + \varepsilon_{t+1}$$

e considerando (3.1), tem-se o modelo auto-regressivo

$$x_{t+p} = f^p(X_t) + \varepsilon_{t+p}, \quad (3.16)$$

onde $E(\varepsilon_{t+p} | X_k, k \leq t) = 0$ (explicação detalhada desse modelo é dada em Yao and Tong, 1994, p.708).

Assim, a previsão de x_t , p passos à frente, pode ser obtida das estimativas locais $f_{T,m,k}^p$ da p -ésima iterada de f no ponto X_t , para a série temporal de comprimento T , na dimensão de imersão m e para o número de pontos k , através das técnicas de regressão não-paramétrica NN ou LWR.

3.3 ESTATÍSTICAS

3.3.1 Erro quadrático médio

Para testar a qualidade dos modelos de previsão (3.16), obtidos com as técnicas de regressão não-paramétrica NN e LWR, são utilizados o erro quadrático médio e a média geométrica (Apêndice C). A série $\{x_t, t=1, \dots, T\}$ é dividida em duas partes. A primeira parte de comprimento T_1 é utilizada para construir as aproximações locais da p -ésima iterada f^p , com as quais serão feitas previsões para os valores da segunda parte da série, de comprimento $T_2=T-T_1$. Os valores das previsões e os valores observados são, então, comparados através dessas estatísticas apontando a qualidade do modelo de previsão.

O erro quadrático médio normalizado EQMN(p) da previsão de x_t , p passos à frente

$$EQMN(p) = \left[\frac{1}{T-p-T_1-1} \sum_{t=T_1+1}^{T-p} (x_{t+p} - \hat{x}_{t+p})^2 \right]^{1/2} / \sigma, \quad (3.17)$$

onde σ é desvio padrão da série temporal, x_{t+p} o valor da série observado p passos à frente e \hat{x}_{t+p} o valor da previsão de x_{t+p} , p passos à frente, é tal que (Farmer and Sidorowich, 1987, p.846):

- 1) se $EQMN(p)=0$, a previsão é perfeita,
- 2) se $EQMN(p)=1$, o desempenho do modelo (3.16) não é superior ao dado pela média da série, isto é, a série tem comportamento semelhante à de ruído branco.

A qualidade das previsões, obtidas com as técnicas de regressão NN e LWR, dependem do:

- 1) número de pontos k na vizinhança de x_t utilizados para construir as estimativas locais da função f^p ,
- 2) número de defasagens no modelo auto-regressivo (3.16), isto é, a dimensão de imersão a ser utilizada para a reconstrução da dinâmica do sistema pelo método das defasagens.

No capítulo 4, técnicas para determinar os valores desses dois importantes parâmetros são discutidas.

3.3.2 Estatística BDS

A validação dos modelos de previsão, obtidos com a técnica de regressão não-paramétrica LWR, pode ser feita segundo Cleveland et al. (1988, p.102), através das mesmas hipóteses que são empregadas para testar modelos obtidos por mínimos quadrados. Neste trabalho, a independência dos resíduos é analisada com o teste BDS (Scheinkman, 1990).

Os testes estatísticos BDS e do terceiro momento incondicional, empregados neste trabalho para analisar não-linearidades em séries temporais, estão resumidos no apêndice A, Testes Estatísticos.

3.4 EXEMPLOS NUMÉRICOS

No quadro 3.1, estão os valores do EQMN(1) para os modelos locais ajustados aos sistemas de Hénon, Hénon6, Hénon22, Lorenz, Lorenzi1.5 e Lorenzi12 com séries de comprimento $T=550$, $T_1=500$ e $T_2=50$. Isto é, para a série de comprimento 550, os 500 primeiros valores foram utilizados para construir aproximações locais (previsões) para os 50 últimos valores observados, com as técnicas de regressão não-paramétrica NN e LWR. O número de pontos k e a dimensão de imersão m foram obtidos com o algoritmo de modelagem apresentado no capítulo 4. No quadro estão, também, os resultados para a série das Manchas Solares⁵ com $T=280$, $T_1=230$ e $T_2=50$.

Os resultados no quadro 3.1 foram obtidos com o esquema de pesos uniforme. As previsões feitas com o esquema de pesos obtidos da função tricúbica (3.13) não apresentaram desempenho superior às previsões com esse esquema mais simples. O quadro 3.2 mostra os resultados obtidos com a regressão não-paramétrica LWR para esses dois esquemas.

QUADRO 3.1-ERRO QUADRÁTICO MÉDIO NORMALIZADO-EQMN(1). REGRESSÃO NÃO-PARAMÉTRICA NN E LWR.

SÉRIE	REGRESSÃO					
	LWR			NN		
	EQMN(1)	m	k	EQMN(1)	m	k
Hénon	0.0020	3	8	0.0374	2	3
Hénon6	0.1331	2	17	0.1511	3	7
Hénon22	0.3607	3	37	0.4078	4	7
Lorenz	0.0136	3	8	0.0810	2	3
Lorenzi1.5	0.0382	3	9	0.0864	2	5
Lorenzi12	0.2349	4	34	0.2324	4	7
M. Solares	0.3986	5	30	0.6135	5	3

$T=T_1+T_2=500+50=550$, $p=1$, $\tau=1$ para as 6 primeiras séries e $T=T_1+T_2=230+50=280$, para a última série. Esquema de pesos uniforme.

QUADRO 3.2—ERRO QUADRÁTICO MÉDIO NORMALIZADO—EQMN(1). REGRESSÃO NÃO-PARAMÉTRICA LWR COM ESQUEMA DE PESOS UNIFORME E TRICÚBICA

SÉRIE	ESQUEMA DE PESOS	
	UNIFORME	TRICÚBICA
Hénon	0.0020	0.0019
Hénon6	0.1331	0.1691
Hénon22	0.3607	0.4434
Lorenz	0.0136	0.0118
Lorenz1.5	0.0382	0.0559
Lorenz12	0.2349	0.2564
M. Solares	0.3986	0.5123

$T=T_1+T_2=500+50=550$, $p=1$, $\tau=1$ para as 6 primeiras séries e $T=T_1+T_2=230+50=280$, para a última série. Valores de m e k dados no quadro 3.1.

Nas figuras 3.1 e 3.2, os valores observados x_{t+1} e as previsões \hat{x}_{t+1} , $t=501, \dots, 549$, para os sistemas dinâmicos de Hénon, Hénon6, Hénon22, Lorenz, Lorenz1.5 e Lorenz12, obtidas com a regressão não-paramétrica LWR, são comparados graficamente (EQMN(1) e demais parâmetros no quadro 3.1).

Para as séries de Hénon e Lorenz, gráficos nas figuras 3.1.a e 3.2.a, respectivamente, não é possível perceber visualmente a diferença, na escala utilizada, entre os valores observados e os valores das previsões. As previsões para séries observadas de sistemas determinísticos caóticos feitas com a regressão não-paramétrica LWR são de ótima qualidade. Tem-se, em geral, $EQMN(p) \rightarrow 0$ com $T \rightarrow \infty$ ((4.9), capítulo 4). Para séries com pequeno comprimento também é possível obter previsões de boa qualidade (EQMN pequeno; definição a seguir), desde que o número de pontos k e a dimensão de imersão m sejam escolhidos corretamente (4.2.1 e 4.2.2, capítulo 4).

Define-se de forma arbitrária, neste trabalho, a partir do

⁵ A série de índices que mede a atividade das Manchas Solares no período 1700–1800 (Tong, 1990, p.470), é não linear, não estacionária e não gaussiana, e serve como referência para métodos estatísticos de modelagem e previsão em séries temporais, que empregam modelos não-lineares. Neste texto, ela serve como exemplo de série do mundo real com pequeno comprimento e com nível de ruído significativo, permitindo esclarecer muitos dos problemas aqui abordados.

FIGURA 3.1. Comparação dos valores observados x_{t+1} com os valores das previsões \hat{x}_{t+1} , $t=501, \dots, 549$, para o sistema dinâmico de Hénon (fig. 3.1.a), Hénon6 (fig. 3.1.b) e Hénon22 (fig 3.1.b). São utilizados 500 valores ($T_1=500$) de cada série para prever os 50 valores seguintes ($T_2=50$), através da regressão não-paramétrica LMR. Para a série Hénon não é possível perceber, na escala utilizada, a diferença entre os valores observados e os valores das previsões. O mesmo ocorre na maior parte dos gráficos correspondentes a Hénon6.

Observado ———, Previsão - - - - -

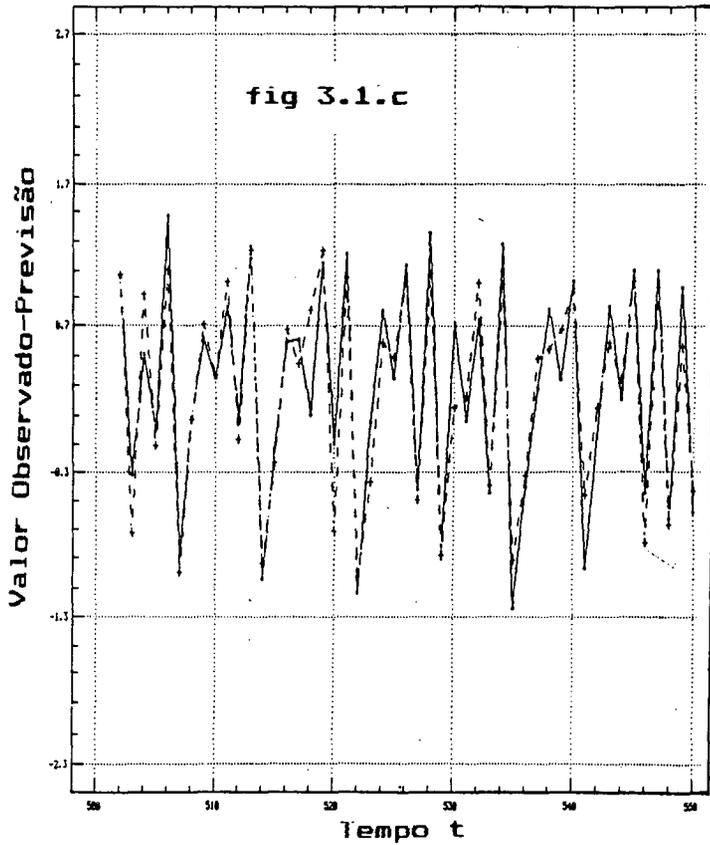
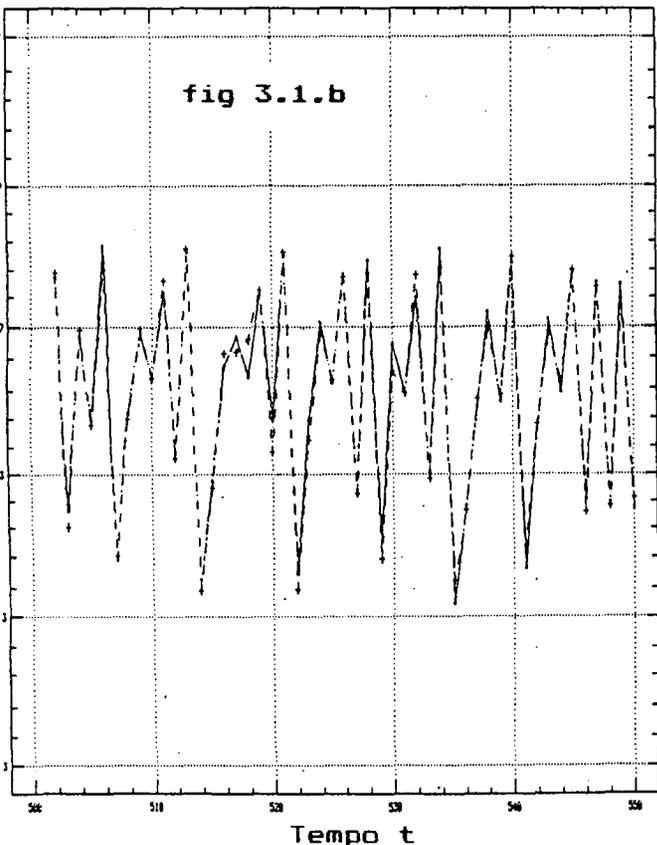
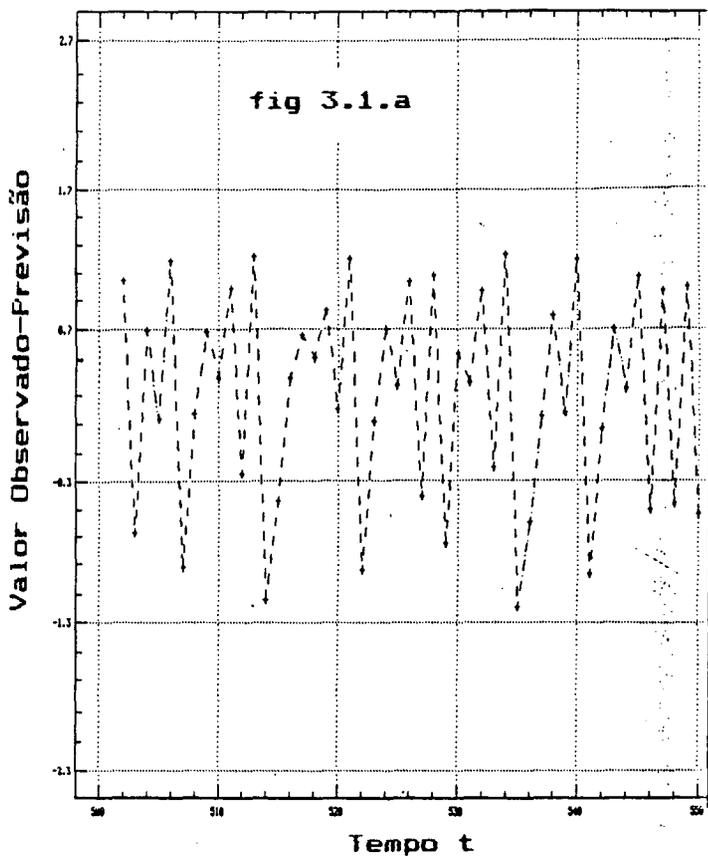


FIGURA 3.2. Comparação dos valores observados x_{t+1} com os valores das previsões \hat{x}_{t+1} , $t=501, \dots, 549$, para o modelo de Lorenz (fig. 3.2.a), Lorenz1.5 (fig. 3.2.b) e Lorenz12 (fig. 3.2.c). São utilizados 500 valores ($T_1=500$) de cada série para prever os 50 valores seguintes ($T_2=50$), através da regressão não-paramétrica LMR. Para a série Lorenz não é possível perceber, na escala utilizada, a diferença entre os valores observados e os valores das previsões. O mesmo ocorre na maior parte dos gráficos correspondentes a Lorenz1.5.

Observado ———, Previsão - - - - -

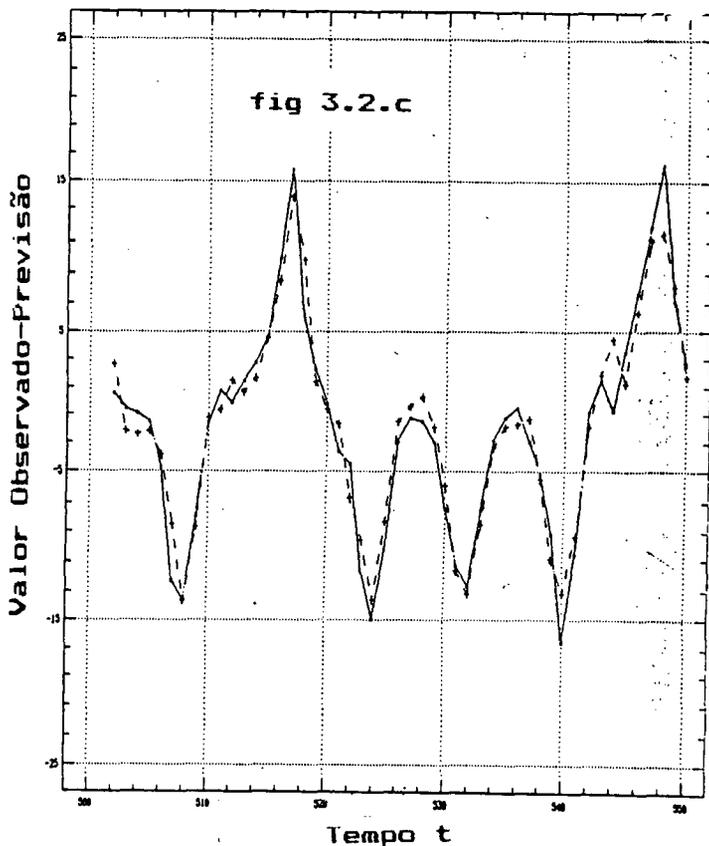
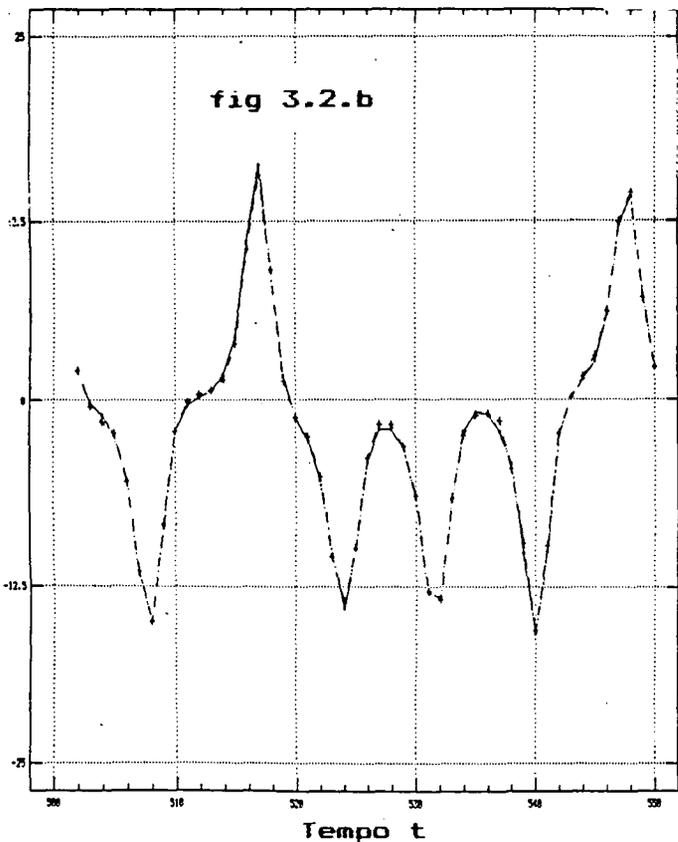
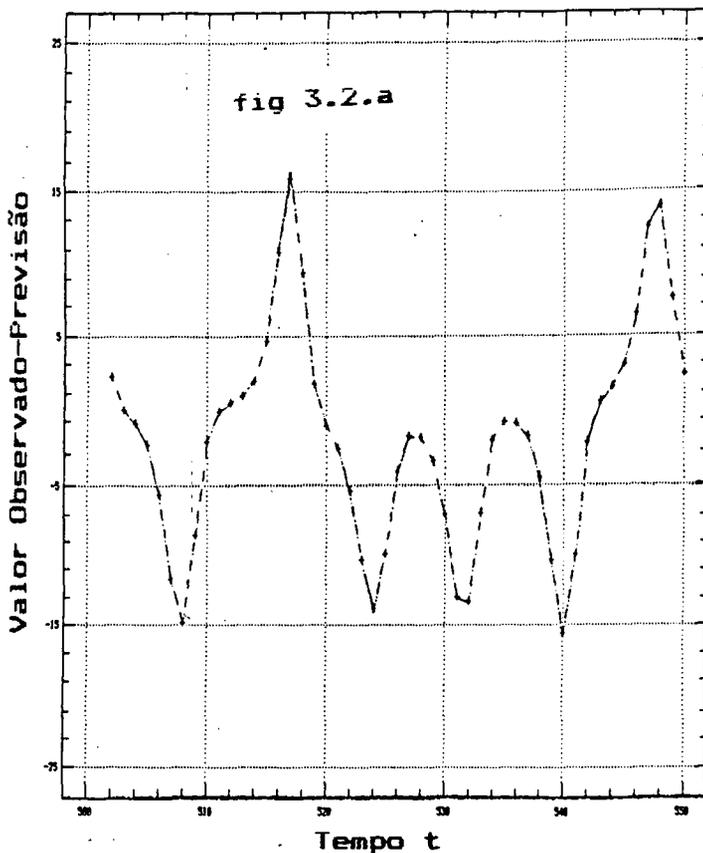


FIGURA 3.3. Comparação dos valores observados x_{t+1} com os valores das previsões \hat{x}_{t+1} , $t=231, \dots, 279$, para a série das Manchas Solares. São utilizados 230 valores ($T_1=230$) para prever os 50 valores seguintes ($T_2=50$), através da regressão não-paramétrica LWR e NN.

Observado ———, Previsão ——— LWR, NN

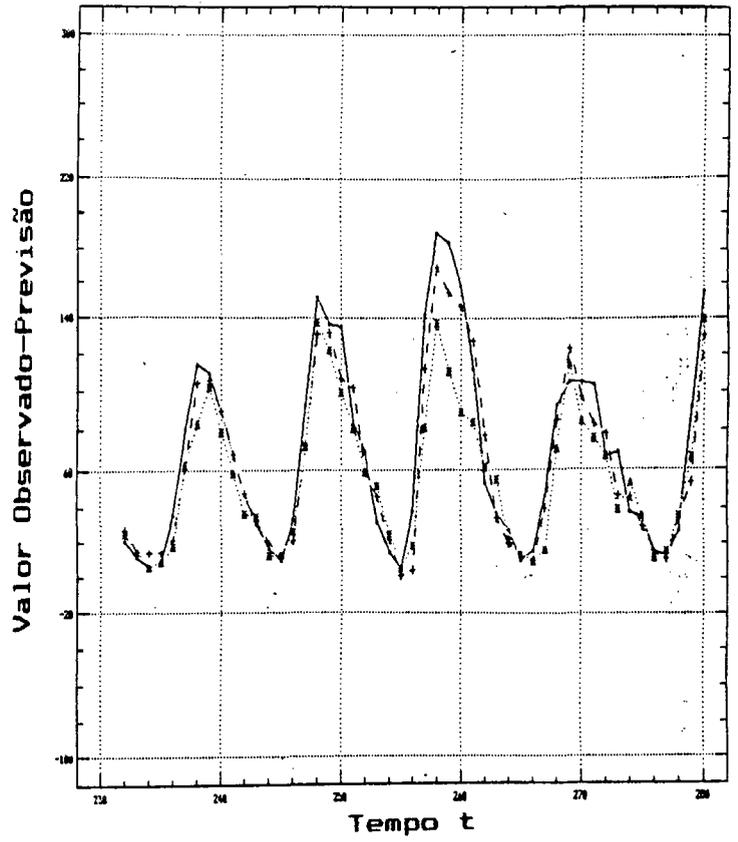
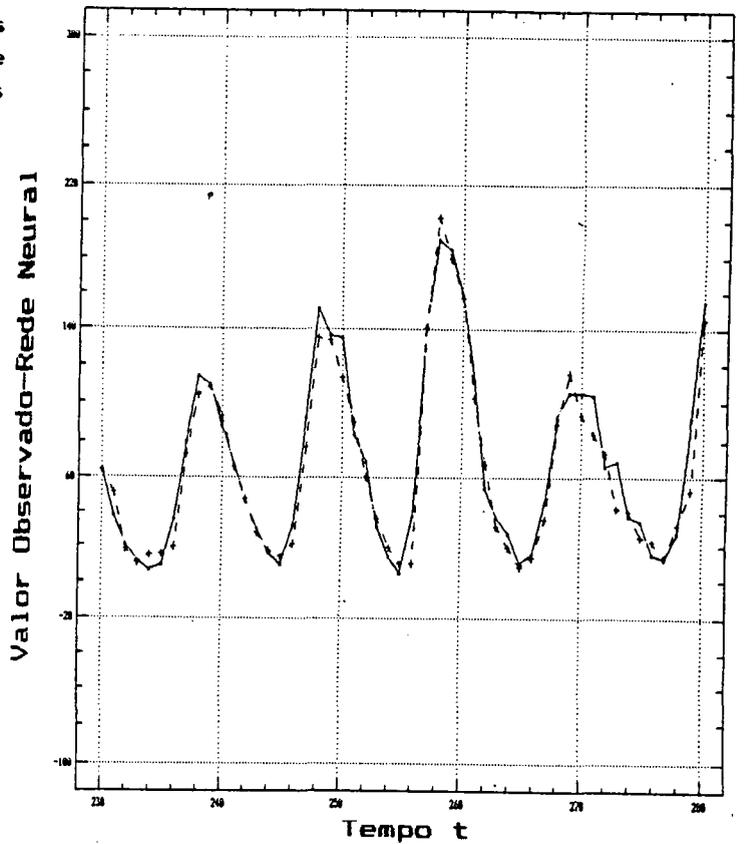


FIGURA 3.8. Comparação dos valores observados x_{t+1} com os valores \hat{x}_{t+1} dados pela rede neural (3.20), ajustada à série das Manchas Solares, $t=229, \dots, 279$.

Observado ———, Rede neural ———



quadro 3.1 e das figuras 3.1 e 3.2, previsão de boa qualidade como sendo aquela que tem $EQMN \leq 0.3$.

Os testes estatísticos para esses modelos locais estão na seção 4.5, do capítulo 4.

A figura 3.3 compara graficamente as previsões \hat{x}_{t+1} , $t=231, \dots, 279$, obtidas com as regressões não-paramétricas NN e LWR, com os valores observados x_t , $t=231, \dots, 279$, para a série das Manchas Solares. Os erros são $EQMN(1)=0.3986$ e $EQMN(1)=0.6135$, respectivamente, para as regressões não-paramétricas LWR e NN.

3.5 REGRESSÃO NÃO-PARAMÉTRICA GLOBAL

Na seção 3.2, as estimativas da p -ésima iterada de f no modelo auto-regressivo (3.16):

$$x_{t+p} = f^p(x_t) + \epsilon_{t+p},$$

são obtidas para cada p através de aproximações locais. Outra possibilidade é obter a aproximação global $f_{T,m}$, a partir da série temporal $\{x_t, t=1, \dots, T\}$ e na dimensão de imersão m , da função $f^1=f$ e usar como aproximação para p -ésima iterada de f (a p -ésima iterada de f dá as previsões de x_t , p passos à frente)

$$f^p_{T,m} = f^{(p-1)}_{T,m} \circ f^1_{T,m}. \quad (3.19)$$

3.5.1 Redes neurais

As aproximações globais $f_{T,m}$ podem ser obtidas através de redes neurais do tipo (Dechert and Gencay, 1992, p.544):

$$f(X) = \sum_{i=1}^N \beta_i \varphi \left(\sum_{j=1}^m w_{ij} x_j + b_i \right), \quad (3.20)$$

as quais segundo Nychka et al. (1992, p.412) podem aproximar funções e suas derivadas com ótima precisão.

O elemento básico das redes, o neurônio, é mostrado esquematicamente na figura 3.4. O neurônio faz a soma ponderada

$$x_1 w_1 + x_2 w_2,$$

dos inputs x_1 e x_2 com os pesos w_1 e w_2 , e a transforma não linearmente, através da função de ativação (transferência) não-linear φ (figura 3.6), produzindo o output

$$O = \psi(w_1x_1 + w_2x_2).$$

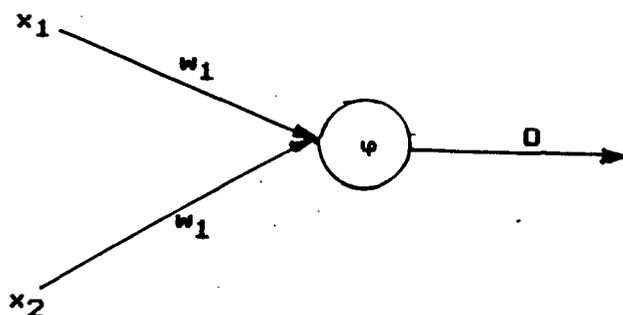


FIGURA 3.4 Representação esquemática do elemento básico da rede neural (3.20), o neurônio. O neurônio faz a soma ponderada dos inputs x_1 e x_2 com os pesos w_1 e w_2 , e a transforma linearmente através da função de ativação (transferência) não-linear ψ (fig. 3.6), produzindo o output $O = \psi(w_1x_1 + w_2x_2)$.

A figura 3.5 ilustra a arquitetura da rede neural (3.20). Os inputs x_1, x_2, \dots, x_m , são recebidos pelas m unidades de input I_j , onde são transformados linearmente e transferidos para as N unidades escondidas U_i . Cada unidade I_j transfere, para a unidade escondida U_i , o input

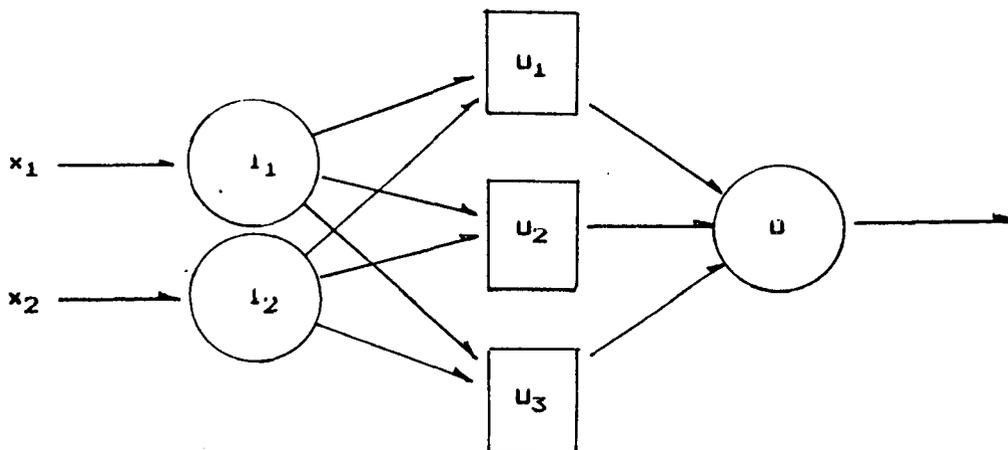


FIGURA 3.5 Arquitetura da rede neural (3.20). As unidades de input I_j recebem os inputs x_j , passando-os para as unidades escondidas U_i . Cada unidade U_i transforma não-linearmente, através da função de ativação ψ , o input total recebido das unidades I_j , passando-os para a unidade de output O . Esta, por sua vez, transforma linearmente os outputs recebidos das unidades escondidas, produzindo o output da rede neural.

$$w_{ij}x_j,$$

onde w_{ij} é o peso do input x_j para a unidade U_i . O input total recebido pela unidade U_i é

$$\sum_{j=1}^m w_{ij}x_j.$$

Cada unidade U_i transforma não-linearmente o input total por ela recebido produzindo o output

$$o_i = \psi\left(\sum_{j=1}^m w_{ij}x_j + b_i\right),$$

onde b_i representa o nível de atividade intrínseco do neurônio. A função de transferência ψ é a mesma para todas as unidades. O output o_i da unidade U_i é passado adiante para a unidade de output, que realiza transformação linear nos o_i produzindo

$$\begin{aligned} f(X) &= \sum_{i=1}^N \beta_i \psi\left(\sum_{j=1}^m w_{ij}x_j + b_i\right) = \\ &= F_{N,m}(X; \beta, W, B), \end{aligned} \quad (3.21)$$

onde $X \in \mathbb{R}^m$ são os inputs, $\beta \in \mathbb{R}^N$ são os pesos para os outputs das unidades escondidas, $W \in \mathbb{R}^{N \times m}$ representa os pesos para os inputs das unidades escondidas e $B \in \mathbb{R}^N$ é o nível de atividade intrínseco dos neurônios.

3.5.2 Função de ativação

A função de ativação ψ (figura 3.6) tem as propriedades:

$$\lim_{x \rightarrow +\infty} \psi(x) = 1 \quad \text{e} \quad \lim_{x \rightarrow -\infty} \psi(x) = 0.$$

As funções $\psi_1, \psi_2 : \mathbb{R} \rightarrow \mathbb{R}$, definidas, respectivamente, por

$$\psi_1(x) = 1/(1+e^{-x})$$

e

$$\psi_2(x) = \tanh(x) = (e^x - e^{-x})/(e^x + e^{-x}),$$

são exemplos de funções muito utilizadas em aplicações das redes neurais (Dechert and Gencay, 1992; Groot and Wurtz, 1991a, b).

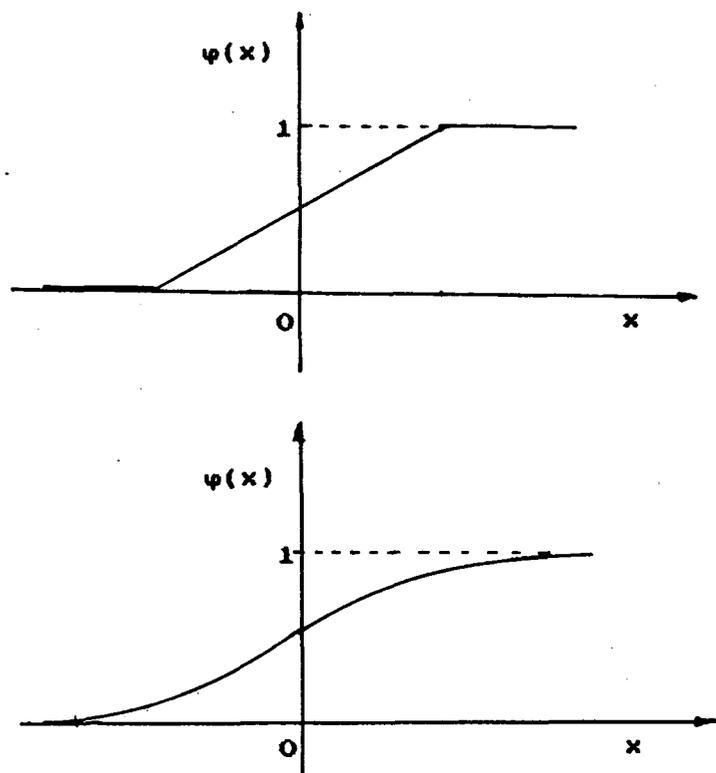


FIGURA 3.6 Exemplos de funções de ativação (transferência) para a rede neural (3.20). A função de ativação é tal que $\psi(x) \rightarrow 1$ com $x \rightarrow \infty$ e $\psi(x) \rightarrow 0$ com $x \rightarrow -\infty$.

Neste trabalho, a função de ativação é (Nychka et al., 1992, p. 411):

$$\psi(x) = x(1 + |x/2|) / (2 + |x| + x^2/2), \quad (3.22)$$

que computacionalmente apresentou melhores resultados.

3.5.3 Estimação da rede neural

Dados $N, m, \{X_t, t=1, \dots, T\}$ e o conjunto de treinamento $Z = \{z_t, t=1, \dots, T\}$, os parâmetros β, W e B devem ser estimados de modo que

$$z_t = f(X_t) = F_{N,m}(X_t; \beta, W, B). \quad (3.23)$$

As redes neurais podem ser estimadas minimizando o critério de desempenho para mínimos quadrados (White, 1989, p. 1005)

$$L(\beta, W, B) = (1/2) \sum_{t=1}^T [(z_t - F_{N,m}(X_t; \beta, W, B))]^2, \quad (3.24)$$

que dá uma medida da qualidade do ajuste do output $F_{N,m}(X; \beta, W, B)$ da rede neural ao conjunto de treinamento Z .

Para o modelo auto-regressivo (3.16), com $p=1$, os elementos z_t do conjunto de treinamento Z são as previsões um passo à frente dos valores x_t da série temporal, isto é,

$$z_t = x_{t+1}$$

e os inputs X_t são as variáveis de estado que reconstroem a dinâmica do sistema na dimensão de imersão m , segundo o teorema de Takes, isto é

$$x_t^m = X_t = (x_t, x_{t-\tau}, \dots, x_{t-(m-1)\tau}),$$

reduzindo o critério de desempenho (3.24) a

$$L(\beta, W, B) = (1/2) \sum_{t=m}^T [x_{t+1} - F_{N,m}(x_t^m; \beta, W, B)]^2. \quad (3.24)$$

Esse problema de mínimos quadrados não-linear pode ser resolvido pelo algoritmo BGFS (Broyden-Fletcher-Goldfarb-Shanno) (Press et al., 1986, p.307). Neste trabalho, a implementação do algoritmo BGFS com derivadas exatas de $L(\beta, W, B)$ e da função de ativação (3.22) mostrou desempenho muito bom, quando comparado com o algoritmo de Levenberg and Marquardt (Press et al. 1986, p.521).

Com a aproximação $f_{T,m}^1 = f_{T,m}$ de f , obtida da série de comprimento T e na dimensão de imersão m , para o modelo auto-regressivo

$$x_{t+1} = f(X_t) + \epsilon_{t+1},$$

a p -ésima iterada de f em (3.16) será obtida de (3.19).

3.5.4 Exemplos Numéricos

Séries determinísticas caóticas e séries não-lineares com ruído

Aos sistemas determinísticos caóticos Logística, Hénon, Mackey-Glass, Ikeda, Lorenz e aos sistemas com ruído Hénon6,

Hénon²², Lorenz 1.5 e Lorenz¹², foram ajustadas redes neurais com a arquitetura (3.20).

Os parâmetros: número de unidades escondidas N , número de inputs (dimensão de imersão) m , e o erro quadrático médio EQM para as séries transformadas para o intervalo $[0,1]$ ⁶, das redes neurais ajustadas, estão no quadro 3.3. O comprimento para todas as séries é $T=200$. Essas redes são empregadas para estimar o expoente de Lyapunov dominante para séries com pequeno comprimento, visando analisar sensibilidade às condições iniciais no sentido introduzido por Nycka et al. (1992), complementando os resultados obtidos com o algoritmo de modelagem no capítulo 4.

A série das Manchas Solares

Na discussão a seguir, pretende-se chamar a atenção para a técnica de diagnosticar caos através da dimensão de correlação em séries com pequeno comprimento e nível de ruído significativo.

A rede neural ajustada à série das Manchas Solares⁷ tem os parâmetros: comprimento $T=280$, número de unidades escondidas $N=5$ e número de inputs $m=5$. O erro quadrático médio para essa rede, com a série transformada para o intervalo $[0,1]$, é $EQM=0.06$.

Na figura 3.7, os valores observados e os valores das previsões um passo à frente, $t=4, \dots, 279$, da série Manchas Solares são comparados graficamente. A figura 3.8 amplia a figura 3.7 para $t=229 \dots, 279$, e pode ser comparada visualmente à figura 3.3, que mostra o mesmo tipo de gráficos para essa parte da série obtidos com as regressões não-paramétrica LWR e NN.

Os coeficientes e os erros padrões estimados para essa rede estão no quadro 3.4.a e 3.4.b (compare-se com (3.20)). O teste estatístico do terceiro momento incondicional (quadro 3.6), que tem o poder de detectar não-linearidades na média (A.2.2, apêndice A), aponta não-linearidades na média dessa série (valor crítico 2.57, a um nível de significância de 1%). O mesmo teste aplicado aos resíduos do modelo (quadro 3.6), dado pela rede neural (3.20),

⁶ A forma funcional da função de ativação (3.22) obriga, para aumentar a eficiência computacional do algoritmo de estimação da rede neural, a transformação da série para o intervalo $[0,1]$.

⁷ Groot and Wurtz (1991a,b) comparam modelos não-lineares clássicos para essa série ao modelo dado por rede neural. Concluem que o desempenho da rede é superior às dos modelos TAR (Tong and Lim, 1980), BL (Priestley, 1981, p.883) e não inferior à do modelo SAR (Priestley, 1981, p.882).

não apresenta estatísticas significativas, mostrando que a rede neural explica as não-linearidades observadas na média dessa série. No quadro 3.5, estão as estatísticas BDS para $\epsilon=0.75$, 1.00 e 1.25 DP., para os dados observados e para os resíduos da rede. As estatísticas para os resíduos não são significativas (valor crítico 2.57, a um nível de significância de 1%), indicando que esses resíduos são iid.

A figura 3.9 apresenta o gráfico da dimensão de correlação para a série das Manchas Solares, estimadas com o algoritmo de Brock-Baeck em (2.7) com $\epsilon=1.0$ DP, e a lei de escala $\nu(m)=m$ (2.1.4, capítulo 2), que vale para processos estocásticos iid. Pode-se observar que, nesse gráfico, $\nu(T,m)$ tende para $\nu^*(m)\approx 3.0$, com m crescente (tende a tornar-se independente de m , com m crescente), comportamento semelhante ao de séries com grande comprimento e com baixo nível de ruído observadas em sistemas determinísticos caóticos. Por outro lado, a discussão acima mostra que o sistema dinâmico dado pela série das Manchas Solares pode ser modelado por modelo estocástico auto-regressivo não-linear.

Observa-se, também, na figura 3.9, o gráfico da dimensão de correlação dos resíduos do modelo não-linear ajustado a essa série, seguindo a lei de escala próxima à $\nu(m)=m$, a qual indica, segundo o algoritmo de Grassberger and Procaccia (seção 2.1.4), que os resíduos do modelo são iid.

Pode-se concluir, a partir da discussão anterior, que a aplicação de algoritmos da dimensão de correlação, ignorando a sua adequação a dada série temporal pode conduzir a resultados duvidosos. Quais são as séries que podem ser investigadas através desses algoritmos? Essa pergunta é respondida no capítulo 4, pelo algoritmo de modelagem. Séries determinísticas caóticas com baixo nível de ruído são modeladas com $EQMN\approx 0.0$. Para a série das Manchas Solares, $EQMN\approx 0.39$.

FIGURA 3.7. Comparação dos valores observados x_{t+1} com os valores \hat{x}_{t+1} dados pela rede neural (3.20) ajustada à série das Manchas Solares, $t=4, \dots, 279$.

Observado ———, Rede neural-----

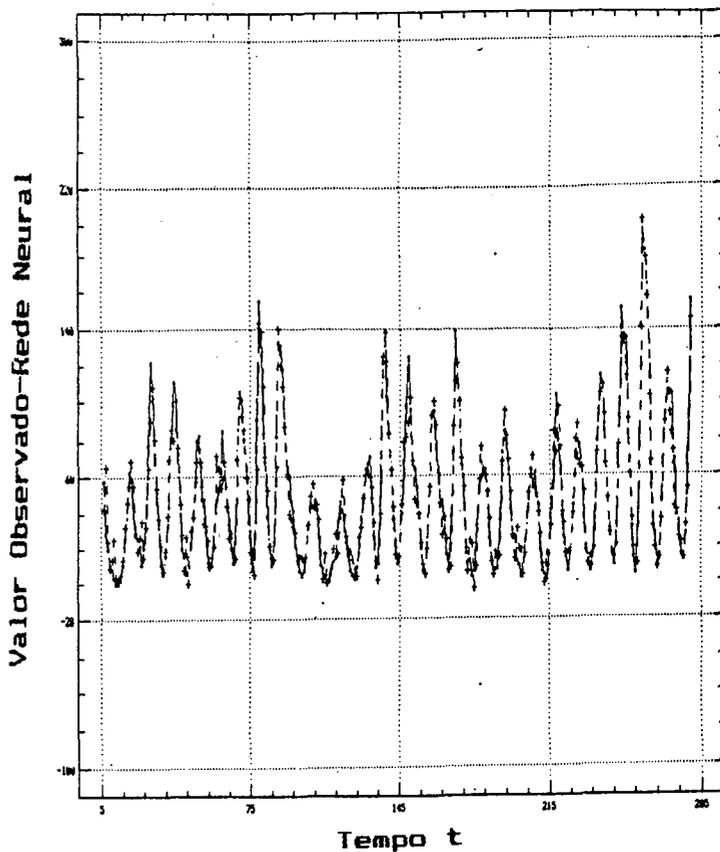
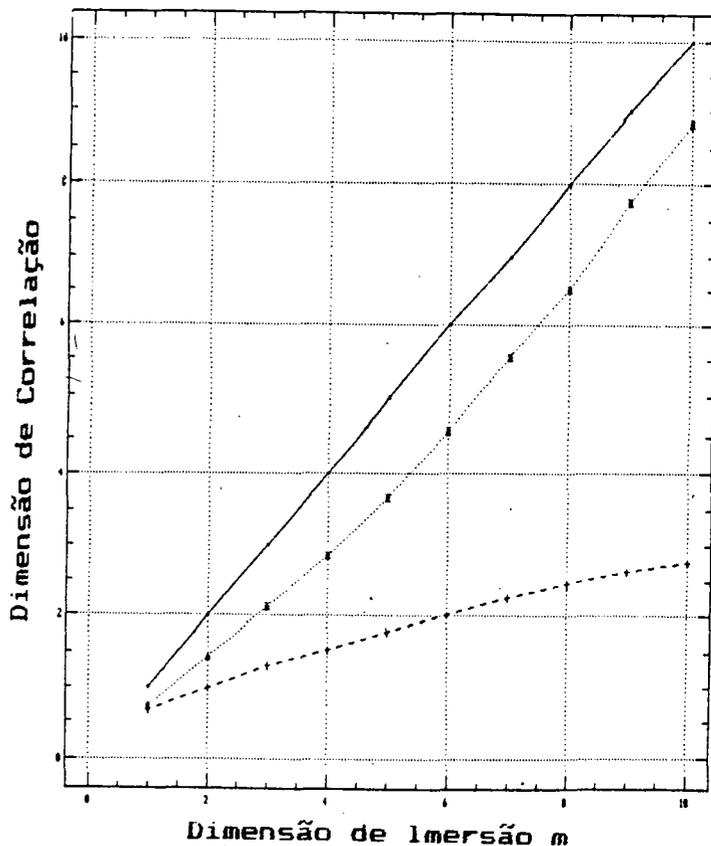


FIGURA 3.9. Dimensão de correlação, estimada através de (2.7 e 2.8, seção 2.2.3), para a série das Manchas Solares e dos resíduos da rede neural (3.20) ajustada à série; dimensão de correlação da lei de escala $\nu(m)=a$ (seção 2.1.4).

M. Solares , Resíduos ----- ,
 $\nu(m)=a$ ———



QUADRO 3.3-REDES NEURAIS. PARÂMETROS N E m, ERRO QUADRÁTICO MÉDIO EQM

SÉRIE	Parâmetros		EQM
	N	m	
Logística	4	1	0.0000
Hénon	5	3	0.0000
Hénon6	5	2	0.0269
Hénon22	5	3	0.0869
M.-Glass	5	4	0.0006
Ikeda	6	4	0.0063
Lorenz	5	3	0.0051
Lorenz1.5	5	3	0.0086
Lorenz12	5	4	0.0414
M. Solares	5	5	0.0606

N=número de unidades escondidas, m=número de inputs=dimensão de imersão, T=200 para as 9 primeiras séries e T=280 para a série das M. Solares. EQM para as séries transformadas para o intervalo [0,1].

QUADRO 3.4.a-COEFICIENTES E ERROS PADRÕES DA REDE NEURAL AJUSTADA À SÉRIE DAS MANCHAS SOLARES(Continua no quadro 3.4.b)

j	Coeficientes				
	w _{1j}	w _{2j}	w _{3j}	w _{4j}	w _{5j}
1	-4.0335 (1.2487)	-2.2951 (1.5178)	-0.7095 (1.0709)	-3.9082 (0.6219)	4.7057 (0.7437)
2	-5.4557 (0.8397)	6.3685 (0.5291)	4.3366 (1.3322)	14.6550 (1.1293)	-7.1624 (0.9263)
3	-12.3659 (2.0227)	7.5773 (2.5855)	16.7365 (0.8965)	19.0939 (3.7997)	-5.6766 (2.1792)
4	-0.4959 (0.3507)	1.1006 (0.5986)	8.2341 (0.3868)	1.6179 (0.7899)	0.6726 (0.4769)
5	-2.4902 (1.8158)	7.1645 (1.5995)	1.7477 (1.1707)	15.6023 (1.3113)	-8.7112 (0.7343)

Os valores entre parênteses são os erros padrões das estimativas dos coeficientes da rede neural.

QUADRO 3.4.b—COEFICIENTES E ERROS PADRÕES DA REDE NEURAL AJUSTADA À SÉRIE DAS MANCHAS SOLARES (Continuação do quadro 3.4.a)

i	Coeficientes	
	β_i	b_i
1	0.4425 (0.0883)	-0.5811 (0.5566)
2	-2.8560 (0.6069)	0.9159 (0.4384)
3	0.6098 (0.1721)	0.5798 (0.8551)
4	1.3250 (0.1688)	0.3349 (0.0922)
5	2.0284 (0.1688)	1.0476 (0.3674)

Os valores entre parênteses são os erros padrões das estimativas dos coeficientes da rede neural.

QUADRO 3.5—ESTATÍSTICA BDS PARA OS RESÍDUOS DA REDE NEURAL AJUSTADA À SÉRIE DAS MANCHAS SOLARES

m	ESTATÍSTICA BDS					
	0.75 DP		1.00 DP		1.25 DP	
	OBS	RES	OBS	RES	OBS	RES
2	34.88	1.21	28.63	1.41	24.55	1.08
3	40.38	0.97	29.80	1.57	24.21	1.42
4	50.72	0.51	33.09	1.23	25.05	1.26
5	64.26	0.33	37.27	0.75	26.48	1.03
6	82.60	0.23	42.41	0.04	28.20	0.51
7	107.60	0.63	48.49	-0.16	30.13	0.51
8	147.60	0.44	58.16	-0.46	33.18	0.29
9	214.79	-0.34	72.76	-0.77	37.94	0.20
10	327.48	-0.23	95.66	-0.80	45.26	0.23

OBS são os valores observados e RES são os resíduos do modelo dado pela rede neural; m é a dimensão de imersão. Valor crítico 2.57, a um nível de significância de 1%.

QUADRO 3.6-TESTE DO TERCEIRO MOMENTO INCONDICIONAL PARA A SÉRIE DAS MANCHAS SOLARES

DEF i,j	SÉRIE	
	OBS	RES
1,1	6.87	0.52
2,1	6.86	-1.10
2,2	6.77	0.12
3,1	7.09	0.27
3,2	6.91	-0.25
3,3	6.93	-0.12
4,1	8.08	0.86
4,2	7.51	0.09
4,3	9.59	-0.33
4,4	7.51	-0.43
5,1	9.59	-0.33
5,2	8.97	1.60
5,3	8.33	0.66
5,4	8.21	-0.30
5,5	8.52	-1.34

OBS são os valores observados e RES são os resíduos da rede neural. DEF indica as defasagens que estão sendo testadas (apêndice A). Valor crítico 2.57, a um nível de significância de 1%.

CAPÍTULO 4

MODELAGEM DE SÉRIES TEMPORAIS NÃO-LINEARES

Os modelos locais de previsão, obtidos com as regressões não-paramétricas do capítulo 3, permitem decidir qual é a modelagem mais adequada para o sistema dinâmico gerador de dada série temporal. É possível decidir qual entre os modelos: determinístico não-linear, estocástico auto-regressivo não-linear ou estocástico auto-regressivo linear, é o mais apropriado. Neste capítulo, são apresentados um algoritmo de modelagem para séries temporais não-lineares, testes estatísticos para os modelos locais(não-lineares) e testes para caos determinístico.

4.1 MODELAGEM

A investigação de caos determinístico de baixa dimensão, em séries temporais com os algoritmos da dimensão de correlação e do expoente de Lyapunov dominante, é desenvolvida na literatura, geralmente, abstraindo-se das análises da adequação dessas técnicas a dada série temporal. Assim, não é surpreendente que os algoritmos da dimensão de correlação identifiquem um processo com raiz unitária como sendo caótico, ou que um ARCH seja identificado pelos algoritmos do expoente de Lyapunov, como um processo que tem sensibilidade às condições iniciais. Quais são as séries que podem ser investigadas com essas técnicas? A resposta está na segunda parte do teorema de Takens, que mostra que sistemas determinísticos caóticos podem ser modelados por modelos determinísticos auto-regressivos não-lineares.

Para decidir qual a modelagem mais adequada para dada série temporal, utiliza-se o algoritmo proposto por Casdagli(1992) para analisar séries determinísticas caóticas. Esse algoritmo é adaptado para as técnicas de regressão não-paramétrica local do capítulo 3 e é utilizado com outra ênfase. As séries com grande comprimento e baixo nível de ruído são substituídas por séries com pequeno comprimento e nível de ruído significativo. A validação dos modelos através do erro quadrático médio normalizado(ou de invariantes dinâmicos e geométricos conhecidos) é substituída por testes estatísticos adequados nos resíduos dos modelos.

4.1.1 O algoritmo de modelagem

Casdagli(1992,p.307) analisa séries determinísticas caóticas com o algoritmo:

Dada a série temporal $\{x_t, t=1, \dots, T\}$,

- 1) Escolhe-se a dimensão de imersão m , o tempo de previsão p e o número k de pontos que serão utilizados em cada vizinhança. Escolhe-se também a defasagem τ ,
- 2) Divide-se a série em duas partes: $\{x_t, t=1, \dots, T_1\}$, que será utilizada para construir as aproximações locais $f_{T,m,k}^p$ da p -ésima iterada de f em

$$x_{t+p} = f^p(x_t) + \epsilon_{t+p}, \quad (4.1)$$

e $\{x_t, t=T_1+1, \dots, T\}$, que será utilizada para testar a qualidade das aproximações locais.

- 3) Escolhe-se o vetor teste

$$X_t = (x_t, x_{t-\tau}, \dots, x_{t-(m-1)\tau})$$

com $t > T_1$, para o qual será feita a previsão p passos à frente.

- 4) Estima-se, utilizando a regressão não-paramétrica linear LWR, a aproximação linear local $f_{T,m,k}^p$ de f^p .
- 5) Estima-se com o modelo em 4)

$$\hat{x}_{t+p} = \hat{x}_{t+p}(T, m, k) = f_{T,m,k}^p(X_t), \quad (4.2)$$

a previsão de x_t , p passos à frente, para o vetor teste X_t e calcula-se o erro

$$e_t(T, m, p, k) = |\hat{x}_{t+p}(T, m, k) - x_{t+p}|. \quad (4.3)$$

- 6) Repetem-se os passos (3)-(5), para todo t no conjunto de teste, e calcula-se o erro quadrático médio normalizado

$$EQMN(T, m, p, k) =$$

$$\left\{ \frac{1}{T-p-T_1-1} \sum_{t=T_1+1}^{T-p} [e_t(T, m, p, k)]^2 \right\}^{1/2} / \sigma. \quad (4.4)$$

Para a regressão não-paramétrica NN, a aproximação linear local em 4) e a previsão de x_t , p passos à frente em 5), são substituídas por

$$\hat{x}_{t+p} = \sum_{i=1}^k w_i x_{i+p}^* \quad (4.5)$$

onde os valores x_{i+p}^* são obtidos dos pontos X_{1+p}^* , com X_1^*, \dots, X_k^* , os k pontos mais próximos de X_t (Por exemplo, suponha-se que X_{23} é o quinto ponto mais próximo de X_t . Então, de $X_{23} = X_5^*$ tem-se $X_{5+p}^* = X_{23+p}$, donde $x_{5+p}^* = x_{23+p}$), e w_i são os respectivos pesos.

4.2 AS CURVAS EQMN(T, m, p, k)

O comportamento das curvas EQMN(T, m, p, k), obtidas a partir das aproximações locais $f_{T, m, k}^P$ de f , em função dos parâmetros T , m , p e k , como apontado por Casdagli(1992, p.307), permite ampla análise do sistema dinâmico gerador da série temporal.

4.2.1 Modelagem determinística vs. modelagem estocástica

Para T , m , p fixos e k variando no intervalo $[K_1, K_2]$, com

$$K_{\min}^1 \leq K_1 \leq k \leq K_2 \leq T_1 - p, \quad (4.6)$$

o comportamento da curva EQMN(T, m, p, k)=EQMN(k) permite decidir qual entre os modelos: determinístico não-linear, estocástico auto-regressivo não-linear e estocástico auto-regressivo linear, é o mais apropriado para modelar o sistema dinâmico do qual foi

¹ O valor mínimo para o extremo inferior K_1 de (4.6) é $k_{\min} \geq m+1$ para a regressão não-paramétrica LWR, isto é, o número mínimo de pontos necessários para dar estabilidade ao algoritmo QR que implementa essa regressão. Neste trabalho adota-se $K_{\min} \approx 2(m+1)$. O extremo superior deve ser escolhido de forma parcimoniosa pois o tempo computacional e a capacidade computacional necessária, crescem rapidamente com k . A implementação do algoritmo de modelagem neste trabalho permite variar o valor de k sem perder os resultados obtidos previamente, isto é, a análise das curvas para $K_2^* > K_2$ ($K_1^* < K_1$) pode ser feita a partir dos resultados obtidos para $k \in [K_1, K_2]$, acrescentando-se os resultados obtidos para $k \in [K_2, K_2^*]$ ($k \in [K_1^*, K_1]$).

observada a série temporal. Para um sistema determinístico caótico, a curva EQMN(k) assume valor mínimo para k pequeno, isto é, o sistema é melhor modelado com k próximo ao menor valor possível para o extremo inferior do intervalo. A partir desse valor de k, a curva EQMN(k) é crescente. Valores de k próximos ao menor valor de (4.6), correspondem à modelagem determinística como a proposta por Farmer and Sidorowich(1987). Para um sistema estocástico auto-regressivo linear, a curva EQMN(k) é decrescente, isto é, o sistema é melhor modelado com k no extremo superior do intervalo. Sistema estocástico auto-regressivo não-linear é modelado com valor de k no interior do intervalo (4.6).

4.2.2 Dimensão do modelo

Com T, p e k fixos, a análise do comportamento da curva EQMN(T,m,p,k)=EQMN(m) permite determinar a dimensão do modelo, isto é, o número de variáveis auto-regressivas no modelo (4.1). Variando m a partir de m=1, a curva EQMN(m) é decrescente com m se aproximando de m^* , m^* a dimensão do modelo. A partir desse valor de m, a curva é crescente. Para os sistemas dinâmicos que não são auto-regressivos, tal comportamento não será observado.

Para um sistema determinístico caótico o erro de previsão EQMN(m) decrescerá para valor próximo a zero com m crescendo para o valor correto da dimensão de imersão m^* (Casdagli,1989,p.345), chamada de dimensão de imersão ideal ou mínima. Se a série temporal tiver nível de ruído significativo tal decréscimo não será observado.

Na dimensão de imersão mínima m^* o processo de reconstrução da dinâmica do sistema não-linear, dada pelo teorema de Takens, é de melhor qualidade. Consequentemente, as estatísticas não-lineares obtidas nessa dimensão de imersão são mais precisas. É importante observar, também, que o número de dados necessários para analisar séries temporais não-lineares, com alguns dos algoritmos neste texto, cresce rapidamente com a dimensão de imersão.

4.2.3 Previsão de curto prazo

Com T, m e k fixos, a curva EQMN(T,m,p,k)=EQMN(p) permite analisar a qualidade das previsões de curto prazo que podem ser feitas para o sistema dinâmico não-linear, a partir da série temporal observada. Para um sistema determinístico caótico, a curva EQMN(p) cresce rapidamente com p, $p=1,2,\dots$ (quanto maior for o expoente de Lyapunov dominante, mais rápido é o crescimento

do erro), mostrando que o horizonte de previsão é limitado a baixos valores de p .

Para os sistemas determinísticos caóticos com baixo nível de ruído, vale a lei de escala (Farmer and Sidorowich, 1987, p.847):

$$EQMN(p) \simeq C e^{h(\alpha+1)p}, \quad (4.7)$$

onde, para a regressão não-paramétrica LWR, $\alpha=1$ e h é a entropia métrica (soma dos expoentes de Lyapunov positivos) do sistema dinâmico associado à série. Para a regressão não-paramétrica NN, $\alpha=0$ e h é expoente de Lyapunov dominante.

De (4.7), tem-se

$$\log EQMN(p) \simeq \log C + h(\alpha+1)p, \quad (4.8)$$

o que permite estimar o maior expoente de Lyapunov ou a entropia métrica h , dependendo do tipo de regressão não-paramétrica utilizada nos modelos de previsão.

4.3 RESULTADOS NUMÉRICOS

Três grupos de sistemas foram analisados com o algoritmo de modelagem em 4.1.1:

- GRUPO I - Sistemas determinísticos caóticos: Logística, Hénon, Ikeda, Lorenz e Mackey-Glass.
- GRUPO II - Sistemas estocásticos: AR(2), ARCH(2), Manchas Solares.
- GRUPO III - Sistemas não-lineares com ruído: Hénon6, Hénon22, Lorenz1.5, Lorenz12.

O grupo I é formado pelos sistemas determinísticos caóticos apresentados no capítulo 1. No grupo II, sistemas estocásticos, os sistemas AR(2) e ARCH(2) são simulados artificialmente, enquanto o sistema das Manchas Solares é dado por uma série do mundo real. Os sistemas do grupo III, sistemas não-lineares com ruído, são obtidos dos sistemas determinísticos caóticos Hénon e Lorenz adicionando-se diferentes níveis de ruído.

As curvas $EQMN(k)$ e $EQMN(m)$ foram estimadas, para os sistemas dos grupos I, II e III, com os parâmetros $p=1$, $\tau=1$, $k \in [20, 420]$, $m=1, \dots, 10$, $T_1=500$, $T_2=200$ e $T=700$ (para a série das Manchas Solares $T_1=230$, $T_2=50$ e $T_3=280$). Na análise em 4.3.1 e 4.3.2, as curvas apresentadas são as curvas com os menores erros (EQMN) para essa especificação dos parâmetros.

4.3.1 Modelagem determinística vs. modelagem estocástica

Na figura 4.1, estão os gráficos das curvas $EQMN(k)$, $k \in [20, 420]$, para os sistemas dos grupos I, II e III.

Para os sistemas dos grupo I, as curvas $EQMN(k)$, gráficos na figura 4.1.a, são estritamente crescentes no intervalo $[20, 420]$. Variando k no intervalo $[2(m+1), 20]$, obtém-se que os menores valores do $EQMN(k)$ são assumidos para $k=2(m+1)$, menor valor possível para o extremo inferior do intervalo, mostrando que a modelagem determinística é a mais apropriada para esses sistemas (o nível de ruído nesses sistemas é o dado pela simulação em Fortran 77, com reais em precisão dupla das correspondentes equações no capítulo 1).

Para os sistemas do grupo II, os gráficos das curvas $EQMN(k)$, $k \in [20, 420]$, estão na figura 4.1.b. A curva correspondente ao ARCH(2), com os valores $EQMN(k) \approx 1.0$, mostra que esse sistema é reconhecido pelo algoritmo de modelagem como ruído (para o algoritmo de modelagem ruído e modelo que não é auto-regressivo são sinônimos). Para os sistemas que não são auto-regressivos, espera-se para a curva $EQMN(k)$ comportamento semelhante ao apresentado pela curva do ARCH(2). A curva $EQMN(k)$, correspondente ao AR(2), decresce de forma monótona no intervalo $[20, 420]$ e assume valor mínimo para k na extremidade superior desse intervalo, mostrando que a modelagem linear é a mais apropriada. Para os sistemas estocásticos auto-regressivos lineares, espera-se para a curva $EQMN(k)$ o mesmo comportamento. Para o sistema das Manchas Solares, figura 4.1.b, a curva $EQMN(k)$ assume valor mínimo em $k=30$, ponto interior do intervalo $[20, 230]$, mostrando que a modelagem através de modelo estocástico auto-regressivo não-linear é a mais apropriada para essa sistema. Observa-se na figura 4.2.a, a qual amplia a figura 4.1.b para $k \in [10, 100]$, o comportamento das curvas $EQMN(k)$ para os sistemas AR(2) e Manchas Solares reforçando, assim, as conclusões anteriores.

Para os sistemas do grupo III, gráficos das curvas $EQMN(k)$ na figuras 4.1.c, $k \in [20, 420]$, e na correspondente ampliação 4.2.b, $k \in [10, 100]$, os valores mínimos são assumidos em pontos interiores do intervalo $[10, 420]$: $k=17$ para Hénon6, $k=37$ para Hénon22 e $k=34$ para Lorenz12, mostrando que a modelagem estocástica auto-regressiva não-linear é a mais apropriada para esses sistemas. A curva $EQMN(k)$ para o sistema Lorenz1.5 assume valor mínimo em $k=10$, indicando para esse sistema, como mais apropriada, a modelagem determinística.

Para os sistemas estocásticos ou determinísticos, apontados pelo algoritmo de modelagem como auto-regressivos não-lineares, o expoente de Lyapunov dominante pode ser utilizado para investigar

a sensibilidade às condições iniciais. No sentido amplo introduzido por Nychka et al.(1992) e Yao and Tong(1994), expoente de Lyapunov positivo em sistemas estocásticos indica sensibilidade às condições iniciais, mostrando que não-linearidades contribuem para a imprevisibilidade do sistema. Para os sistemas determinísticos, expoente de Lyapunov positivo indica sensibilidade às condições iniciais, o que é traço característico da presença de caos. No quadro 4.2, estão os expoentes de Lyapunov dominantes para os modelos apontados como auto-regressivos pelo algoritmo de modelagem. As estimativas foram feitas com séries de comprimento $T=200$ (para a série das Manchas Solares, $T=280$), aproximando a função f (F em 2.15, capítulo 2) geradora da dinâmica e o Jacobiano ($J_t=DF(X_t)$) em 2.16, capítulo 2) através de redes neurais, com os parâmetros dados no quadro 3.3(Algoritmo Global, capítulo 2).

Os valores positivos, estimados para os expoentes de Lyapunov, mostram que os modelos auto-regressivos analisados apresentam sensibilidade às condições iniciais. Na terminologia de Yao and Tong(1994), são sistemas determinísticos e estocásticos caóticos.

Para a modelagem de séries observadas do mundo real, em geral com ruído, é importante a colocação de Casdagli(1992, p.306):

Se os efeitos do ruído são pequenos e a dimensão D do atrator é baixa, então, com séries de comprimento moderado, boas aproximações da função f podem ser feitas com as técnicas de regressão não-paramétricas em dimensão $m>D$. Por outro lado, se a dimensão D do atrator é grande, o comprimento da série observada pode não ser suficiente para aproximar f com modelo determinístico não-linear em dimensão de imersão $m>D$. É possível que o sistema dinâmico seja melhor modelado por modelo estocástico em dimensão $m<D$.

Essa colocação alerta para o fato que sistemas determinísticos caóticos de alta dimensão(1.2.7, capítulo 1) podem ser identificados através do algoritmo de modelagem como sendo ruído e, também, como processos estocásticos auto-regressivos, sem que o sejam. Assim, o ruído pode estar escondendo um sistema caótico de alta dimensão. É nesse sentido que não existe, operacionalmente, diferença entre caos de alta dimensão e ruído.

QUADRO 4.1—EXPOENTES DE LYAPUNOV DOMINANTE PARA OS SISTEMAS AUTO-REGRESSIVOS

SISTEMA	E. LYAPUNOV
Logística	0.6932
Hénon	0.4067
Hénon6	0.4173
Hénon22	0.2655
Ikeda	0.3369
Lorenz	0.2292
Lorenz1.5	0.1329
Lorenz12	0.0574
M.-Glass	0.1209
M. Solares	0.0272

Expoentes de Lyapunov estimados de séries com comprimento $T=200$, através do algoritmo GLOBAL com redes neurais. Para a série das M. Solares, $T=280$.

4.3.2 Dimensão do modelo

A figura 4.3 apresenta os gráficos das curvas $EQMN(m)$, $m=1, \dots, 10$, para os sistemas dos grupos I, II e III.

As curvas $EQMN(m)$, $m=1, \dots, 10$, para os sistemas do grupo I, gráficos na figura 4.3.a, indicam dimensão de imersão mínima (ou ideal) m^* igual a 3 para Hénon, a 4 para Ikeda, a 3 para Lorenz e a 4 para Mackey-Glass (para a Logística, gráfico truncado na figura 4.3.a, $m^*=1$). Uma pesquisa mais detalhada, no entorno dos valores de m e k obtidos da análise das curvas $EQMN$, e com k variando no intervalo $[2(m+1), 420]$, conduz aos valores ótimos m^* e k^* , que minimizam $EQMN(m, k)$, apresentados no quadro 4.2.

Pode-se observar, nos gráficos das curvas $EQMN(m)$ na figura 4.3.a, uma característica marcante dos sistemas determinísticos caóticos com baixo nível de ruído, que é a queda brusca do $EQMN(m)$ com m se aproximando da dimensão de imersão mínima m^* e o crescimento do $EQMN(m)$ para valores de m superiores a m^* (para Ikeda por exemplo, com os valores ótimos dos parâmetros k e m , quadro 4.2, tem-se $EQMN(5) \approx 10EQMN(4)$). Essa característica, em particular, dá fundamento às críticas feitas aos métodos de diagnosticar caos em séries temporais do mundo real que trabalham com dimensão de imersão crescente. De forma geral, para séries com comprimento limitado, o erro $EQMN(m)$ cresce com a dimensão de imersão m . Para contornar esse problema, os métodos que trabalham com dimensão de imersão crescente exigem séries com grande comprimento, para tornar $EQMN(m)$ suficientemente pequeno para

todo m .

Para os sistemas determinísticos caóticos com baixo nível de ruído, vale a lei de escala (Casdagli, 1992, p.309)

$$EQMN(T) \approx C(k/T_1)^{\alpha/D}, \quad (4.9)$$

onde $\alpha=1$ para regressão não-paramétrica NN, $\alpha=2$ para regressão não-paramétrica LWR e D é a dimensão do atrator (a mais apropriada é a dimensão de informação; para atratores com baixa dimensão não existem diferenças significativas entre as várias definições de dimensão (Smith, 1992, p.330)). A lei de escala (4.9) mostra que $EQMN(T) \rightarrow 0$ com $T \rightarrow \infty$, se $k/T_1 \rightarrow 0$.

A figura 4.5 apresenta os gráficos das curvas $EQMN(T)$, obtidas com $T_1=400, 600, \dots, 1200$, $T_2=300$, $T=T_1+T_2$ e os demais parâmetros com os valores dados no quadro 4.2, para os sistemas de Lorenz e Hénon. Observa-se na figura 4.5 a curva, correspondente a Lorenz, seguindo a lei de escala (4.9) (a curva log-log é aproximadamente linear; (C.3), apêndice C). No apêndice C, explica-se a razão pela qual a lei de escala (4.9) não é seguida com fidelidade pelo sistema de Hénon, com $T=400, 600, \dots, 1200$.

Outra prática, encontrada na literatura, é escolher a dimensão de imersão m como o menor inteiro maior que a dimensão de correlação (ou outra dimensão qualquer). Por exemplo, em Farmer and Sidorowich (1987, p.845): "If the attractor is of dimension D , a minimal requirement is that $m \geq D$." Para Ikeda, por exemplo, a dimensão de correlação é $\nu \approx 1.67 < 2$, donde $m=2$ pode ser escolhida, mas a dimensão de imersão ideal é $m^*=4$. Os resultados corretos obtidos com as dimensões de imersão dadas pelo critério $m \geq D$, devem ser creditados às séries com grande comprimento utilizadas pelos pesquisadores de caos em Física. Para séries temporais do mundo real, em geral com pequeno comprimento, o conhecimento da dimensão de imersão mínima é, como apontado por Sugihara and May (1990, p.738) e Cheng and Tong (1992, p.429), muito importante para evitar perda de informações dadas pela série.

Para os sistemas do grupo II, os gráficos das curvas $EQMN(m)$ estão na figura 4.3.b. Para o sistema ARCH(2), a curva $EQMN(m)$ é crescente com m , $m=1, \dots, 10$. Adotou-se, então, para esse sistema, $m^*=1$. Para os sistemas estocásticos AR(2) e Manchas Solares², as curvas $EQMN(m)$ indicam, respectivamente, $m^*=2$ e $m^*=5$.

² Groot and Wurtz (1991a, b) ajustam rede neural à série das M. Solares. Variando o número m de variáveis auto-regressivas (número m de unidades de inputs) concluem que os resultados não são melhores para $m \geq 5$. A discussão acima confirma essa conclusão, obtida pelos autores por tentativa e erro (Groot and Wurtz, 1991b, p.185).

QUADRO 4.2-VALORES ÓTIMOS DOS PARÂMETROS m E k QUE MINIMIZAM O ERRO QUADRÁTICO MÉDIO $EQMN(m,k)$

GRUPO	Parâmetros		EQMN(m^* , k^*)
	m^*	k^*	
I			
Logística	1	4	0.0001
Hénon	3	8	0.0025
Ikeda	4	10	0.0068
Lorenz	3	8	0.0155
M.-Glass	4	10	0.0062
II			
AR(2)	2	450	0.2810
ARCH(2)	1	-	≈ 1.0
M. Solares	5	30	0.3986
III			
Hénon6	2	17	0.1535
Hénon22	3	37	0.4049
Lorenz1.5	3	9	0.0389
Lorenz12	4	34	0.2314

Valores ótimos obtidos com $T=T_1+T_2=500+200$, $p=1$, $\tau=1, \dots, 10$, $k=[2(m+1), 20]$, regressão não-paramétrica LWR com esquema de pesos uniforme. Para a série das M. Solares, $T=T_1+T_2=230+50=280$.

Para os sistemas do grupo III, os gráficos das curvas $EQMN(m)$, $m=1, \dots, 10$, estão na figura 4.3.c. As curvas indicam $m^*=2$ para Hénon6, $m^*=3$ para Hénon22, $m^*=3$ para Lorenz1.5 e $m^*=4$ para Lorenz12. Observe-se que a dimensão de imersão ideal para o sistema determinístico puro e para o correspondente sistema com ruído não são, obrigatoriamente, coincidentes.

4.3.4 Previsão de curto prazo

As curvas $EQMN(T, m, p, k) = EQMN(p)$ com T , m , e k fixos, permitem analisar as previsões de curto prazo que podem ser feitas para os sistemas dinâmicos, a partir das séries temporais observadas.

Na figura 4.4, estão os gráficos das curvas $EQMN(p)$, $p=1, \dots, 10$, para os sistemas do grupo I. O traço característico dessas curvas é o rápido crescimento do $EQMN(p)$ com o tempo de previsão p . A taxa de crescimento do $EQMN(p)$ é função do expoente

de Lyapunov dominante. Quanto maior o expoente de Lyapunov, maior é a taxa de crescimento do EQMN(p). Dessas curvas, pode-se estimar através de (4.8) a entropia métrica ou o expoente de Lyapunov dominante, dependendo do tipo de regressão empregada nos modelos de previsão.

No apêndice C, Entropia Métrica e Expoente de Lyapunov, a lei de escala (4.7) é analisada para os sistemas dinâmicos do grupo I, utilizando-se séries com comprimento $T=T_1+T_2=1000+500$. Os valores dos expoentes de Lyapunov, estimados das curvas EQMN(p) através da regressão (4.8), são comparados com os respectivos valores obtidos com os algoritmos do capítulo 2. Além disso, outros problemas relevantes para os modelos de previsão locais de séries determinísticas caóticas são destacados nesse apêndice.

Do quadro 4.3.a, que apresenta os erros de previsão EQMN(p), $p=1, \dots, 10$, para os sistemas do grupo I, pode-se concluir que é possível fazer previsões de boa qualidade ($\text{EQMN}(p) \leq 0.3$; definição em 3.4, capítulo 3) até 4 passos à frente ($p=4$) para 200 valores das séries a partir dos 500 primeiros valores observados para esses sistemas. Em particular, para Mackey-Glass é possível fazer previsões 10 passos à frente ($p=10$) com $\text{EQMN}(p) < 0.23$.

Para os sistemas do grupo II, os erros de previsão EQMN(p) estão no quadro 4.3.b. Duas características de processos estocásticos podem ser observadas a partir desses valores:

1) para processos estocásticos lineares auto-regressivos o EQMN(p) cresce lentamente com o tempo de previsão. Para o AR(2) tem-se $\text{EQMN}(1)=0.28$ e $\text{EQMN}(10)=0.41$;

2) para processos reconhecidos pelo algoritmo como ruído (processos que não são auto-regressivos) deve-se esperar $\text{EQMN}(p)=1.0$ para todo p. Para o ARCH(2) tem-se $\text{EQMN}(1)=1.09$ e $\text{EQMN}(10)=1.11$.

Para os sistemas do grupo III, sistemas não-lineares com ruído, os valores EQMN(p), $p=1, \dots, 10$, estão no quadro 4.3.b. Os valores EQMN(p) mostram que o ruído adicionado ao sistema determinístico caótico faz com que os erros de previsão EQMN(p) sejam substancialmente maiores que os dos correspondentes sistemas determinísticos, como era de se esperar.

Dois problemas interessantes podem ser investigados nesses sistemas:

1) a influência da sensibilidade às condições iniciais (expoente de Lyapunov positivo, quadro 4.1) no erro de previsão, isto é, a decomposição do erro de previsão em duas

partes: uma correspondente ao ruído e outra correspondente à sensibilidade às condições iniciais;

2) a determinação do nível de ruído que pode ser adicionado em dado sistema determinístico caótico sem que ele deixe de apresentar sua característica básica, isto é, sua sensibilidade às condições iniciais.

No quadro 4.3.b, pode-se observar que é possível fazer previsões de boa qualidade 2 passos à frente ($p=2$) para Hénon6 ($EQMN(1)=0.15$, $EQMN(2)=0.25$), 1 passo à frente para Lorenzi2 ($EQMN(1)=0.23$), e 3 passos à frente ($p=3$) para Lorenzi.5 ($EQMN(1)=0.04$, $EQMN(2)=0.08$, $EQMN(3)=0.22$). A comparação dos valores $EQMN(p)$ para Lorenzi.5, com os correspondentes erros do sistema determinístico, mostra que, apesar do baixo nível de ruído nessa série, a qualidade das previsões para o sistema com ruído é reduzida drasticamente em relação ao sistema determinístico.

4.4 DIAGNOSTICANDO CAOS EM SÉRIES TEMPORAIS

As curvas $EQMN(k)$ e $EQMN(m)$ dão poderoso teste para diagnosticar caos determinístico em séries temporais. A curva $EQMN(k)$ indica a modelagem mais apropriada para série. A curva $EQMN(m)$ indica a dimensão de imersão ideal. No caso de apontarem modelagem determinística com $EQMN(m^*) \approx 0.0$, a série temporal pode ser modelada por modelo determinístico não-linear (modelo auto-regressivo não-linear com nível de ruído próximo de zero). Nesses modelos o caos determinístico pode explodir.

Para os sistemas do grupo I e para Lorenzi.5 no grupo III as curvas $EQMN(k)$ indicam modelagem determinística não-linear como a mais apropriada, isto é, os valores mínimos dessas curvas são assumidos em valores próximos ou iguais ao menor valor possível para a extremidade inferior do intervalo (4.6). As curvas $EQMN(m)$ assumem valores mínimos $EQMN(m^*) \approx 0.0$ nas dimensões de imersão ideal m^* , dados no quadro 4.2. Assim, esses modelos são identificados corretamente por esse teste de caos como modelos determinísticos não-lineares. A confirmação da presença de caos nesses modelos pode ser feita estimando-se o expoente de Lyapunov dominante, valores no quadro 4.1, com os algoritmos do capítulo 2 (ou estimando-se das curvas $EQMN(p)$, através da relação (4.8), a entropia métrica ou o expoente de Lyapunov dominante). Para séries com pequeno comprimento o algoritmo GLOBAL é o mais indicado.

FIGURA 4.1. Gráficos das curvas EQMN(k), $k=20, \dots, 420$, para os sistemas dinâmicos do grupo I (Hénon, Ikeda, Lorenz e Mackey-Glass) (fig. 4.1.a), grupo II (AR(2), ARCH(2) e Manchas Solares) (fig. 4.1.b) e grupo III (Hénon6, Hénon22, Lorenz1.5 e Lorenz12) (fig. 4.1.c), obtidas através da regressão não-paramétrica LWR com os parâmetros, $p=1, \tau=1, T_1=500, T_2=200, T=700, m=1, \dots, 10$. As curvas apresentadas são as curvas com o menor erro quadrático médio normalizado.

Grupo I

Hénon ———, Ikeda - - - - ,
Lorenz , M.-Glass - - - - .

Grupo II

AR(2) ———, ARCH(2) - - - - , M. Solares

Grupo III

Hénon6 , Hénon22 - - - - ,
Lorenz1.5 ———, Lorenz12 - - - - .

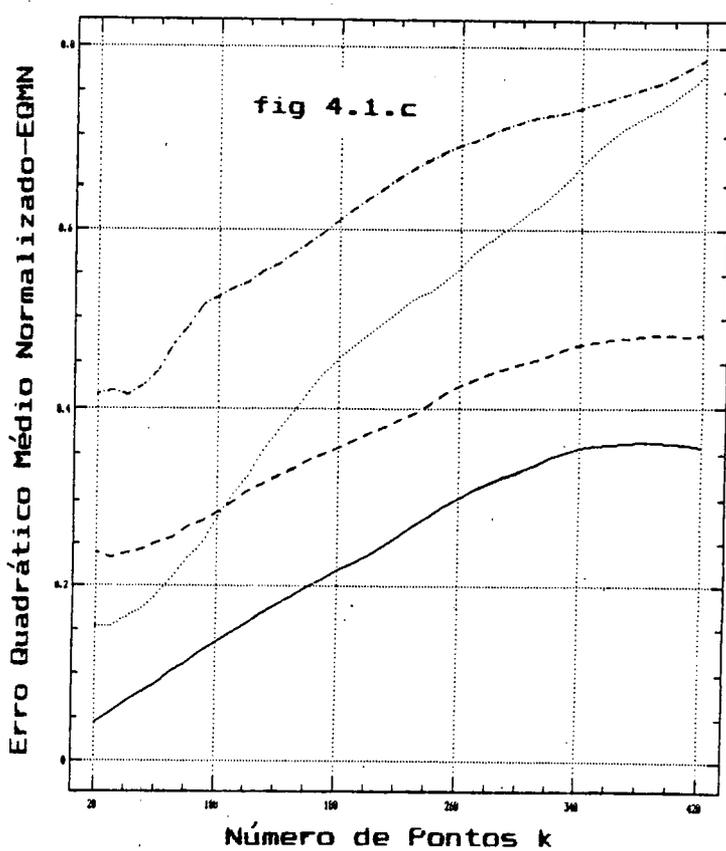
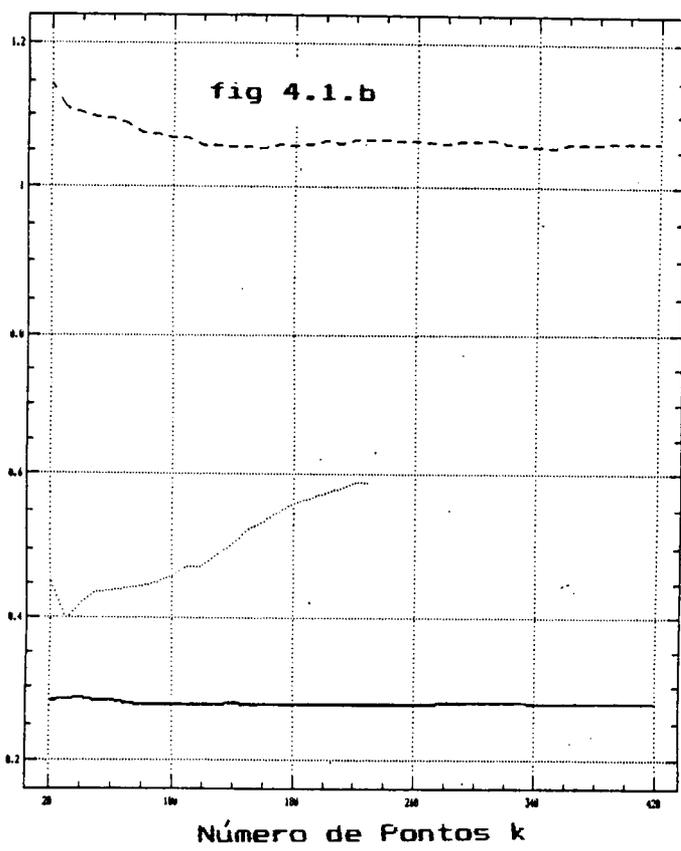
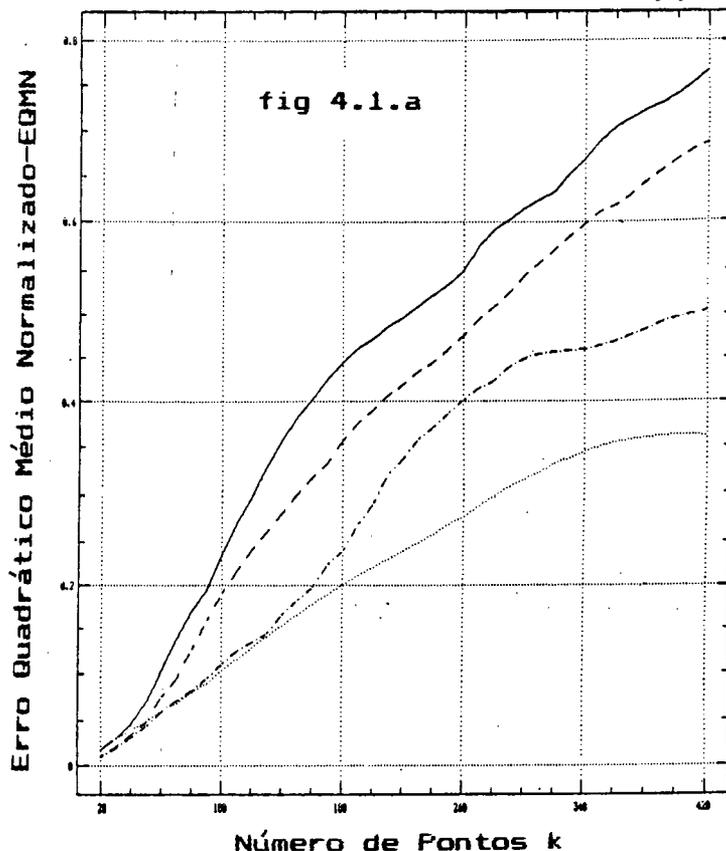


FIGURA 4.2. A figura 4.2 amplia algumas das curvas $EQMN(k)$ na figura 4.1. Na figura 4.2.a tem-se as curvas $EQMN(k)$ para os sistemas dinâmicos AR(2) e Manchas Solares do grupo II, enquanto na figura 4.2.b tem-se as curvas para os sistemas do grupo III (Hénon6, Hénon22, Lorenz1.5 e Lorenz12), $k=10, \dots, 100$.

Grupo II

AR(2) ———, M. Solares - - - - -

Grupo III

Hénon6 , Hénon22 - - - - - ,

Lorenz1.5 ———, Lorenz12 - - - - -

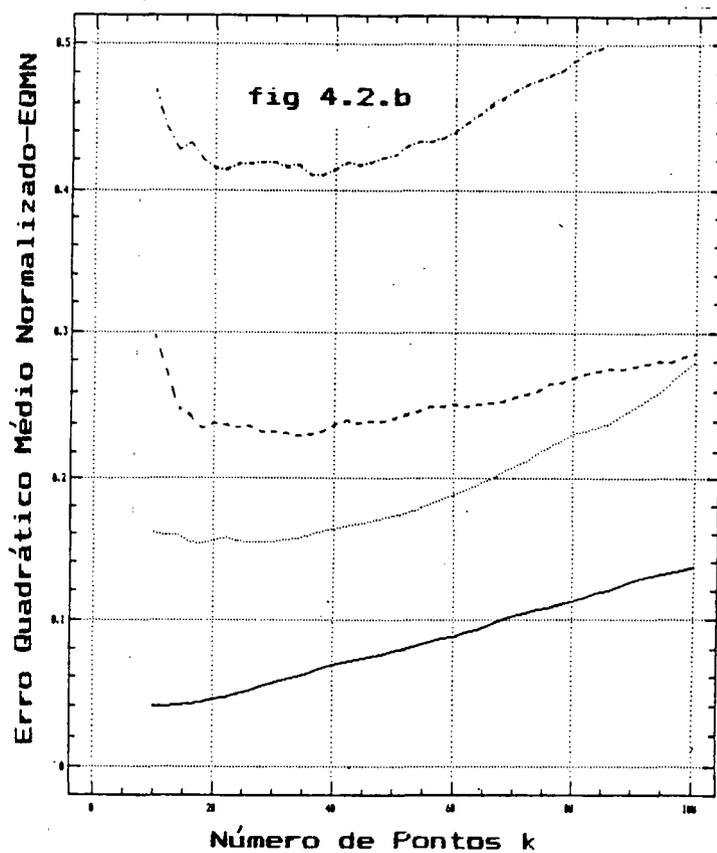
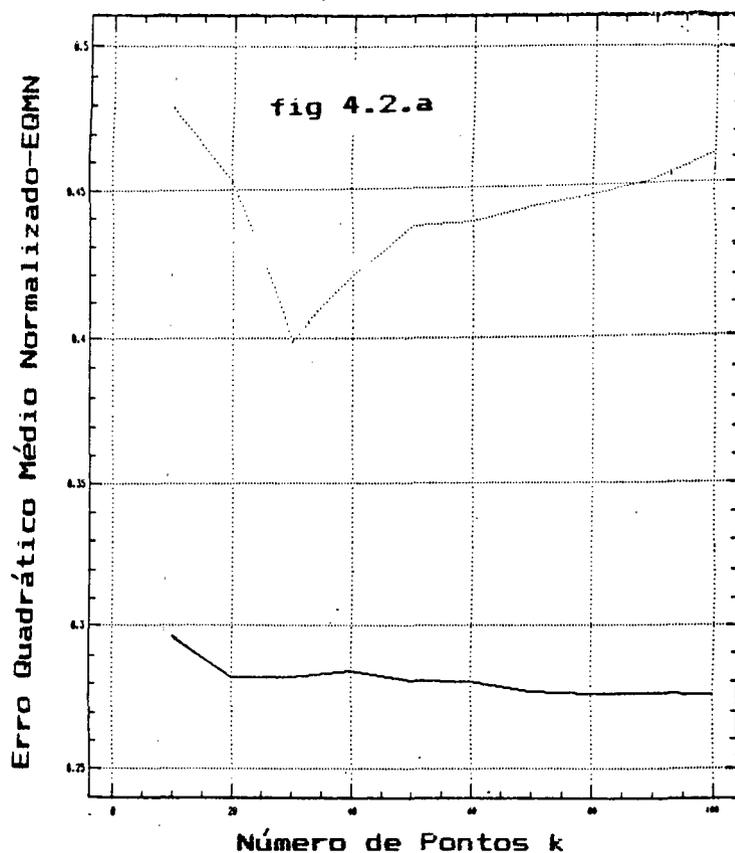


FIGURA 4.3. Gráficos das curvas EQMN(m), $m=1, \dots, 10$, para os sistemas dinâmicos do grupo I (Logística, Hénon, Ikeda, Lorenz e Mackey-Glass) (fig. 4.3.a), grupo II (AR(2), ARCH(2) e Manchas Solares) (fig. 4.3.b) e grupo III (Hénon6, Hénon22, Lorenz1.5 e Lorenz12) (fig. 4.3.c), obtidas através da regressão não-paramétrica LWR com os parâmetros, $p=1$, $\tau=1$, $T_1=500$, $T_2=200$, $T=700$, $k=20, \dots, 420$. As curvas apresentadas são as curvas com o menor erro quadrático médio normalizado.

Grupo I

Logística ———, Hénon ———, Ikeda ,
Lorenz ———, M.-Glass ———

Grupo II

AR(2) ———, ARCH(2) ———, M. Solares

Grupo III

Hénon6 ———, Hénon22 ———,
Lorenz1.5, Lorenz12

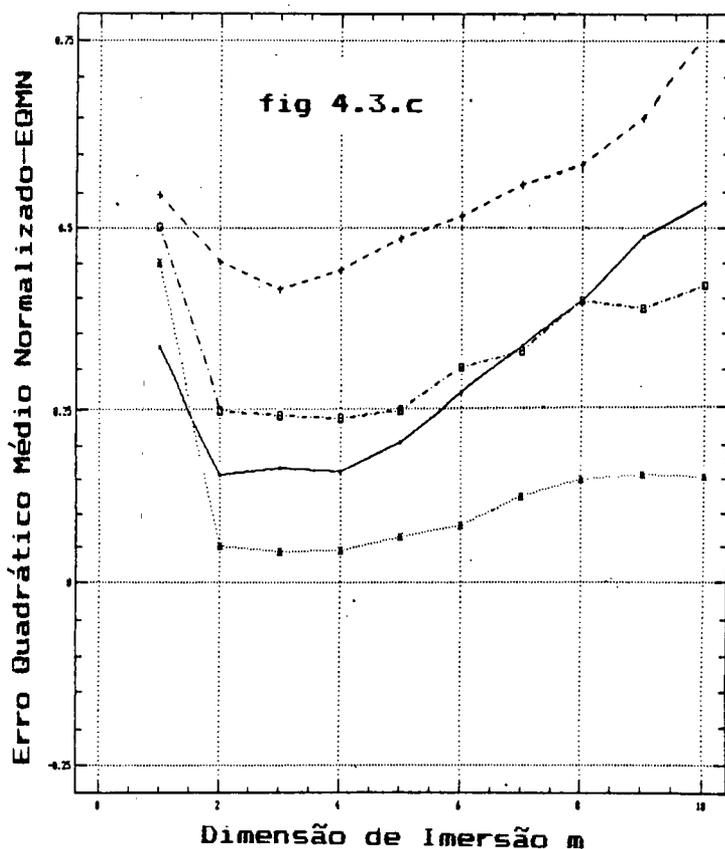
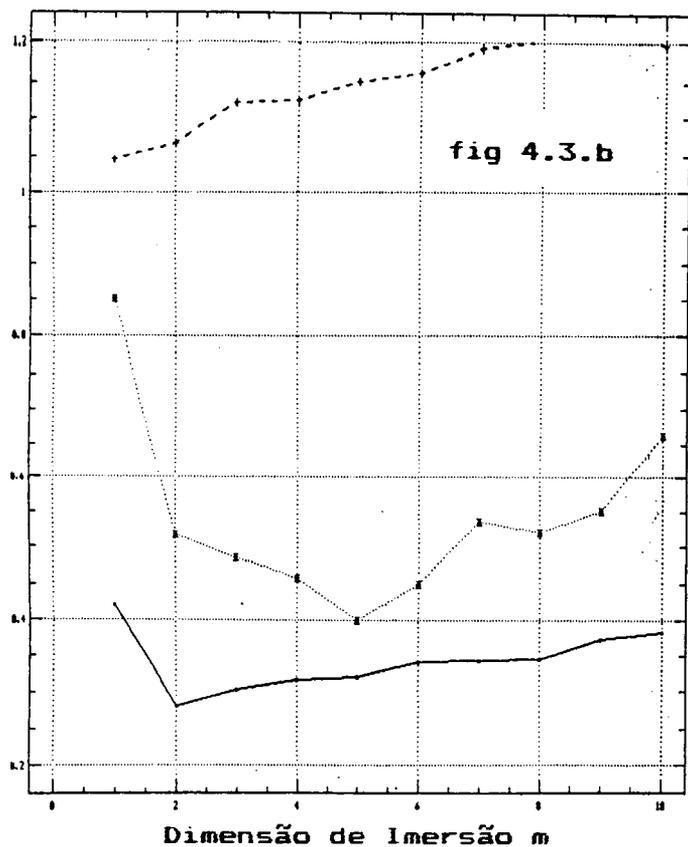
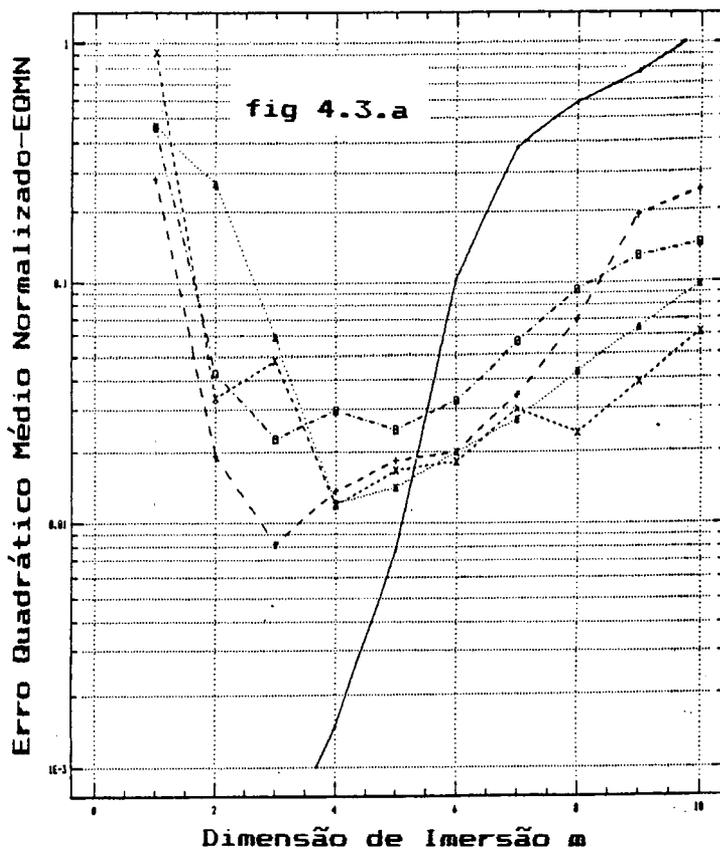


FIGURA 4.4. Gráficos das curvas EQMN(p), $p=1, \dots, 10$, para os sistemas dinâmicos do grupo I (Logística, Hénon, Ikeda, Lorenz e Mackey-Glass), obtidas através da regressão não-paramétrica LMR com os parâmetros, $p=1$, $\tau=1$, $T_1=500$, $T_2=200$, $T=700$, m e k com os valores dados no quadro 4.2, seção 4.3.2.

Logística ———, Hénon - - - - , Ikeda ,
Lorenz , Mackey-Glass - - - - -

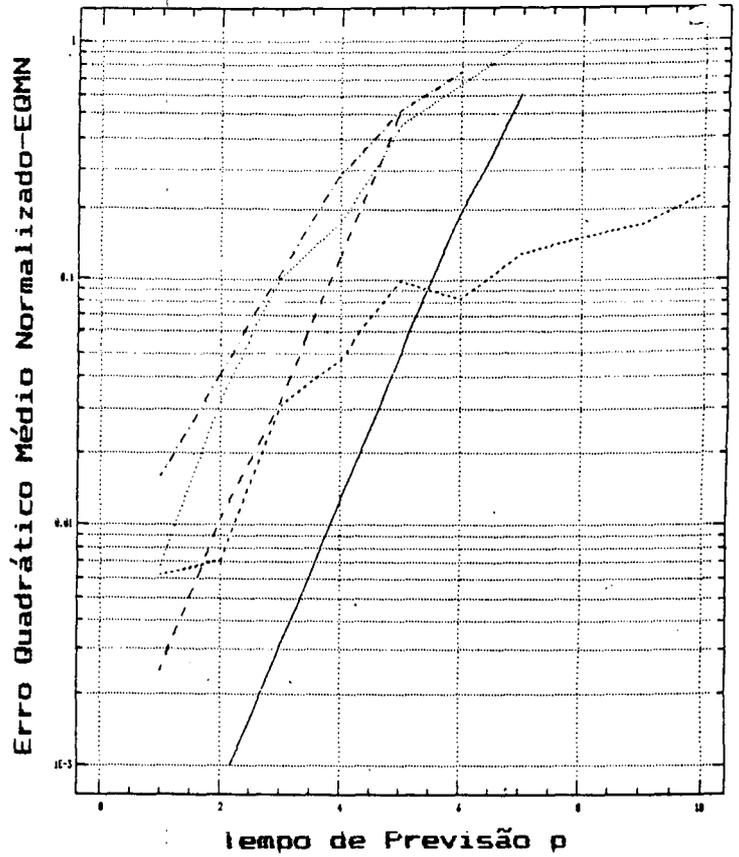
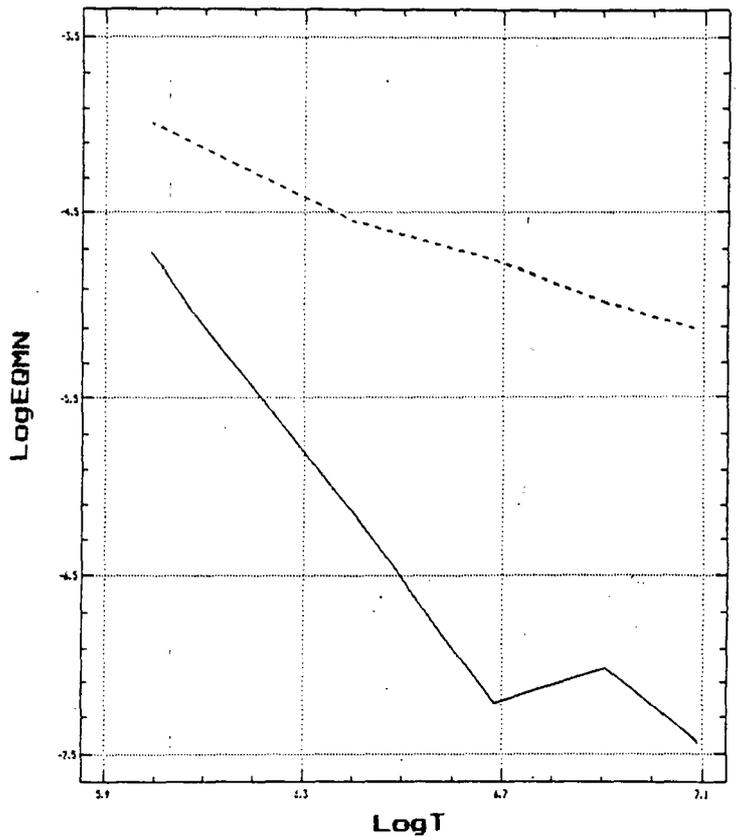


FIGURA 4.5 Gráficos das curvas EQMN(T), $T_1=400, \dots, 1200$, $T_2=300$, $p=1$, $\tau=1$, valores de m e k dados no quadro 4.2, seção 4.3.2, para os sistemas dinâmicos de Hénon e Lorenz.

Hénon ———, Lorenz - - - - -



QUADRO 4.3.a-EQMN(p) PARA OS SISTEMAS DO GRUPO I: SISTEMAS DETERMINÍSTICOS CAÓTICOS

p	SISTEMA DETERMINÍSTICO				
	Logística	Hénon	Ikeda	Lorenz	M.-Glass
1	0.0001	0.0025	0.0068	0.0155	0.0062
2	0.0008	0.0106	0.0324	0.0416	0.0071
3	0.0031	0.0322	0.1012	0.1066	0.0312
4	0.0126	0.1265	0.1763	0.2900	0.0476
5	0.0490	0.5178	0.4524	0.5168	0.0997
6	0.1935	*	0.6225	0.7506	0.0833
7	0.6029	*	0.9951	0.7559	0.1293
8	*	*	*	*	0.1508
9	*	*	*	*	0.1728
10	*	*	*	*	0.2305

EQMN(p) estimado com $T=T_1+T_2=500+200$, $p=1$, $\tau=1$, m e k dados no quadro 4.2, regressão não-paramétrica LWR com esquema de pesos uniforme. Para a série das M. Solares, $T=T_1+T_2=230+50=280$.

QUADRO 4.3.b-EQMN(p) PARA OS SISTEMAS DOS GRUPOS II E III: SISTEMAS ESTOCÁSTICOS E SISTEMAS NÃO-LINEARES COM RUÍDO

p	SISTEMA					
	NÃO-LINEAR COM RUÍDO				ESTOCÁSTICO	
	Hénon6	Hénon22	Lorenz1.5	Lorenz12	AR(2)	ARCH(2)
1	0.1546	0.4192	0.0443	0.2349	0.2810	1.0972
2	0.2535	0.6182	0.0894	0.3489	0.2933	1.1305
3	0.4026	0.8485	0.2230	0.5720	0.3484	1.1315
4	0.6394	0.9733	0.4908	0.7791	0.3608	1.0971
5	0.9731	*	0.6836	0.8500	0.3787	1.0704
6	*	*	0.7162	*	0.3847	1.0740
7	*	*	*	*	0.3954	1.0703
8	*	*	*	*	0.3972	1.0742
9	*	*	*	*	0.3974	1.0799
10	*	*	*	*	0.4151	1.1101

EQMN(p) estimado com $T=T_1+T_2=500+200$, $p=1$, $\tau=1$, m e k dados no quadro 4.2, regressão não-paramétrica LWR com esquema de pesos uniforme. Para a série das M. Solares, $T=T_1+T_2=230+50=280$.

No capítulo 2, os gráficos $\text{Log}C(\epsilon)$ vs. $\text{Log} \epsilon$ são obtidos para os sistemas de Hénon6, Hénon22, Lorenzi1.5 e Lorenzi12, empregando-se séries com comprimento $T=1500$. Esses gráficos são utilizados para diagnosticar caos pelo algoritmo de Grassberger-Procaccia. Escolhe-se a parte linear do gráfico e por regressão linear estima-se a dimensão de correlação. Os gráficos nas figuras 2.3.b, 2.3.c, 2.4.b e 2.4.c, mostram que o efeito do ruído é reduzir a parte linear desses gráficos, na parte mais necessária ao algoritmo ($\epsilon \rightarrow 0$), dificultando ou mesmo inviabilizando as estimativas da dimensão de correlação. Em particular, as partes lineares dos gráficos correspondentes ao modelo de Lorenzi1.5 são reduzidas drasticamente.

A discussão acima mostra um caminho alternativo ao dado pelo excelente algoritmo de Grassberger-Procaccia para séries com grande comprimento e baixo nível de ruído, para investigar caos determinístico em séries com pequeno comprimento.

4.5 TESTES ESTATÍSTICOS

Para analisar a adequação dos modelos locais ajustados aos sistemas Hénon, Hénon6, Hénon22, Lorenz, Lorenzi1.5 e Lorenzi12, são utilizados o teste do terceiro momento incondicional e o teste BDS. A escolha dos testes estatísticos foi baseada nas propriedades que o teste do terceiro momento tem para detectar não-linearidades na média, enquanto o teste BDS é muito sensível a desvios da distribuição iid. Os testes estatísticos estão resumidos no apêndice A, Testes Estatísticos.

4.5.1 Sistemas não-lineares com ruído

Os quadros 4.4 e 4.5 apresentam os testes estatísticos para os modelos locais (não-lineares) de Hénon6 e Hénon22, enquanto os quadros 4.5 e 4.6 apresentam os testes para os modelos locais de Lorenzi1.5 e Lorenzi12. Para esses modelos, os testes do terceiro momento incondicional, quadros 4.4 e 4.6, e BDS, quadros 4.5 e 4.7, aplicados aos resíduos, não apresentam estatísticas significativas (valor crítico 2.57, a um nível de significância de 1%). Assim, os modelos locais (não-lineares) para esses sistemas dinâmicos conseguem capturar as não-linearidades apontadas pelos testes estatísticos aplicados aos dados brutos, quadros 4.4 a 4.7.

4.5.2 Sistemas determinísticos caóticos

Para os modelos ajustados aos sistemas determinísticos caóticos Hénon e Lorenz, o teste do terceiro momento não apresenta estatísticas significativas, quadro 4.8, mostrando que os modelos locais explicam o comportamento não-linear desses sistemas, enquanto o teste BDS mostra que os resíduos não são iid., quadro 4.9.

Observando-se que a regressão não-paramétrica LWR é eficiente para modelos estocásticos lineares com k suficientemente grande e que o valor de k utilizado no modelo de Hénon($k=8$) e Lorenz($k=8$) é o menor valor possível permitido pelo algoritmo QR, pode-se conjecturar que o teste BDS está apontando dependência linear nos resíduos desses modelos.

Filtrando os resíduos do modelo local para Hénon por um AR(7), e do modelo local para Lorenz por um AR(3), tem-se os resultados apresentados no quadro 4.9, para o teste BDS aplicado aos resíduos desses modelos lineares.

Pelas recomendações de Cao and Tsay(1992,p. S171), de que as estatísticas BDS devem ser enfatizadas para $m=2,\dots,5$, pode-se considerar que os resíduos dos modelos não-lineares de Hénon e Lorenz filtrados pelos modelos lineares AR(7) e AR(3), respectivamente, são iid., pois as estatísticas BDS(1.0 DP), nessas dimensões de imersão, não são significativas, conforme mostram os resultados, para $m \leq 5$, no quadro 4.9(valor crítico 2.57, a um nível de significância de 1%).

Assim, os testes do terceiro momento e BDS validam os modelos locais(não-lineares) obtidos para os sistemas de Hénon e Lorenz através da regressão não-paramétrica LWR.

QUADRO 4.4—TESTE DO TERCEIRO MOMENTO PARA OS MODELOS LOCAIS DE HÉNON6 E HÉNON22

DEFASAGEM i, j	Hénon6		Hénon22	
	OBS	RES	OBS	RES
1,1	-5.34	1.33	-5.07	0.96
2,1	2.38	-0.78	2.32	1.22
2,2	-0.48	0.33	-0.07	0.58
3,1	-4.34	-0.12	-4.37	0.72
3,2	4.00	0.17	3.51	-0.07
3,3	0.05	0.58	-0.43	1.32
4,1	0.92	1.77	1.08	-0.42
4,2	-1.65	-0.20	-1.38	-0.34
4,3	0.53	0.05	0.53	-0.70
4,4	3.61	0.26	2.86	1.21
5,1	-0.50	-1.20	-0.70	-0.91
5,2	0.71	-0.40	0.84	-1.24
5,3	-1.02	-1.04	-0.91	-0.97
5,4	-0.98	0.94	-0.90	0.97
5,5	1.31	0.76	0.62	1.21

Na coluna OBS, tem-se as estatísticas para os valores observados da série ($t=501, \dots, 700$), enquanto que na coluna RES tem-se as estatísticas para os resíduos do modelo local (não-linear) ajustado à série ($t=501, \dots, 700$). Valor crítico 2.57, a um nível de significância de 1%.

QUADRO 4.5—ESTATÍSTICA BDS(1.0 DP) PARA OS MODELOS LOCAIS DE HÉNON6 E HÉNON22

m	Hénon6		Hénon22	
	OBS	RES	OBS	RES
2	21.64	-0.00	22.96	-0.57
3	21.79	0.82	23.86	0.64
4	23.81	0.68	27.15	1.01
5	24.50	0.61	29.00	0.77
6	27.88	0.47	33.79	0.43
7	32.86	0.37	39.83	-0.09
8	39.29	0.18	48.41	-0.42
9	48.54	-0.15	59.76	-0.84
10	61.43	-0.26	74.19	-0.69

Na coluna OBS, tem-se as estatísticas para os valores observados da série ($t=501, \dots, 700$), enquanto que na coluna RES, tem-se as estatísticas para os resíduos do modelo local (não-linear) ajustado à série ($t=501, \dots, 700$). Valor crítico 2.57, a um nível de significância de 1%.

QUADRO 4.6-TESTE DO TERCEIRO MOMENTO INCONDICIONAL PARA OS MODELOS LOCAIS DE LORENZ1.5 E LORENZ12

DEF i,j	Lorenz1.5		Lorenz12	
	OBS	RES	OBS	RES
1,1	2.26	1.40	2.14	1.31
2,1	2.95	2.35	2.84	0.65
2,2	3.52	0.11	3.46	0.52
3,1	4.23	0.71	4.07	0.34
3,2	4.67	1.02	4.74	-0.93
3,3	5.49	-0.28	5.46	1.39
4,1	5.25	-0.93	4.89	0.29
4,2	5.61	-0.59	5.31	-0.28
4,3	5.64	-0.15	5.38	-1.05
4,4	4.07	-0.79	3.64	1.01
5,1	4.79	-0.46	4.49	-0.51
5,2	5.34	-0.19	4.84	-1.14
5,3	4.80	-0.17	4.09	-0.27
5,4	2.07	-2.06	1.95	0.11
5,5	0.02	-1.51	0.13	-0.55

Na coluna OBS, tem-se as estatísticas para os valores observados da série($t=501, \dots, 700$), enquanto que na coluna RES tem-se as estatísticas para os resíduos do modelo local(não-linear) ajustado à série($t= 501, \dots, 700$). Valor crítico 2.57, a um nível de significância de 1%.

QUADRO 4.7- ESTATÍSTICA BDS(1.0 DP) PARA OS MODELOS LOCAIS DE LORENZ1.5 E LORENZ12

m	Lorenz1.5		Lorenz12	
	OBS	RES	OBS	RES
2	20.76	2.07	20.72	1.79
3	20.96	1.41	20.76	1.90
4	24.65	1.37	23.95	2.29
5	32.21	0.98	30.55	2.36
6	44.73	0.79	41.67	2.52
7	62.46	0.91	57.04	2.48
8	87.83	0.84	79.54	2.23
9	123.66	1.20	111.37	1.55
10	173.40	1.54	155.06	1.03

Na coluna OBS, tem-se as estatísticas para os valores observados da série($t=501, \dots, 700$), enquanto que na coluna RES tem-se as estatísticas para os resíduos do modelo local(não-linear) ajustado à série($t= 501, \dots, 700$). Valor crítico 2.57, a um nível de significância de 1%.

QUADRO 4.8—TESTE DO TERCEIRO MOMENTO INCONDICIONAL PARA OS MODELOS LOCAIS DE HÉNON E LORENZ

DEF i,j	Hénon		Lorenz	
	OBS	RES	OBS	RES
1,1	-5.38	0.81	2.27	0.35
2,1	2.36	-0.21	2.95	-1.65
2,2	-0.65	-1.63	3.50	0.50
3,1	-4.22	-0.98	4.22	0.84
3,2	4.17	-1.21	4.67	1.10
3,3	0.31	0.07	5.51	0.03
4,1	0.83	0.87	5.25	0.74
4,2	-1.75	2.52	5.62	-0.20
4,3	0.54	-1.41	5.67	2.09
4,4	3.83	-1.55	4.08	-0.71
5,1	-0.42	0.43	4.77	0.18
5,2	0.60	0.33	5.36	-1.20
5,3	-1.08	-1.14	4.84	1.86
5,4	-1.00	-0.66	2.11	0.18
5,5	1.56	-1.40	0.05	-2.36

Na coluna OBS, tem-se as estatísticas para os valores observados da série($t=501, \dots, 700$), enquanto que na coluna RES tem-se as estatísticas para os resíduos do modelo local(não-linear) ajustado à série($t=501, \dots, 700$). Valor crítico 2.57, a um nível de significância de 1%.

QUADRO 4.9—ESTATÍSTICA BDS(1.0 DP) PARA OS RESÍDUOS DOS MODELOS LOCAIS DE HÉNON E LORENZ FILTRADOS POR MODELO LINEAR

m	Hénon		Lorenz	
	RES1	RES2	RES1	RES2
2	3.57	0.28	3.59	0.58
3	3.15	0.56	4.08	1.04
4	2.89	0.81	3.98	1.47
5	2.45	1.05	3.49	1.90
6	2.24	1.31	3.03	2.36
7	2.02	1.57	2.70	2.87
8	2.10	1.87	2.60	3.37
9	2.18	2.11	2.61	3.86
10	2.22	2.35	2.62	4.42

RES1 são os resíduos do modelo local(não-linear) ajustado à série($t=501, \dots, 700$), enquanto que RES2 são os resíduos RES1 filtrados por modelo linear. Valor crítico 2.57, a um nível de significância de 1%.

CAPÍTULO 5

NÃO-LINEARIDADES EM SÉRIES FINANCEIRAS BRASILEIRAS

O algoritmo de modelagem para séries não-lineares no capítulo 4 é aplicado aos retornos das séries financeiras IBOVESPA, BRASFUND, VALE e WHM, buscando-se evidências de variáveis auto-regressivas nos sistemas dinâmicos geradores dessas séries. Testes estatísticos não-lineares e modelos não-lineares na variância, também, são utilizados para detectar e explicar não-linearidades nessas séries financeiras.

5.1 MODELAGEM

As curvas $EQMN(k)$ e $EQMN(m)$ ((4.2.1) e (4.2.2), capítulo 4) foram estimadas para os retornos das séries financeiras (a palavra retornos será subtendida no restante deste capítulo, isto é, série financeira deve ser entendida como a série dos retornos da correspondente série financeira) IBOVESPA, BRASFUND, WHM e VALE, com $\tau=1$, $m=1, \dots, 10$, $p=1$, $k \in [20, 460]$, e com os comprimentos T_1 , T_2 e T dados a seguir:

SÉRIE	Comprimento			Período
	T_1	T_2	$T=T_1+T_2$	
IBOVESPA	1298	300	1598	03/01/86 a 15/04/92
BRASFUND	1279	300	1579	04/05/88 a 30/06/94
WHM	1007	300	1307	02/01/86 a 07/01/91
VALE	1007	300	1307	02/01/86 a 07/01/91

O extremo inferior $k=20$, do intervalo de variação de k , foi escolhido em função do valor mínimo necessário para assegurar a estabilidade numérica do algoritmo ($k \geq 2(m+1) = 2k_{\min}$). O extremo superior leva em consideração o tempo computacional¹, o qual cresce rapidamente com o número de pontos k . Essas escolhas mostraram-se satisfatórias para analisar as quatro séries.

As curvas $EQMN(k)$ e $EQMN(m)$, nas seções 5.1.1 e 5.1.2, são as curvas com os menores erros para a especificação dos

parâmetros acima.

5.1.1 Modelagem linear vs. modelagem não-linear

Para a série IBOVESPA, tem-se $EQMN(k) \approx 1.0$, $k \in [20, 460]$, mostrando que o algoritmo de modelagem interpreta essa série como ruído.

Para as séries WHM e VALE, gráficos na figura 5.1, as curvas $EQMN(k)$ são decrescentes no intervalo $[20, 460]$ (aproximadamente constantes para $k \geq 300$), com essas curvas assumindo valores mínimos na extremidade superior do intervalo $[20, 460]$. Assim, o algoritmo de modelagem indica a modelagem auto-regressiva linear como a mais adequada para essas duas séries. Para a série BRASFUND, o comportamento da curva $EQMN(k)$ é mais complexo que os dois anteriores. A curva $EQMN(k)$, gráfico na figura 5.1, apresenta queda brusca para $k \in [20, 80]$, e continua a decrescer de forma mais lenta para $k \in [80, 220]$, atingindo valor mínimo em $k=220$ (ponto interior do intervalo $[20, 460]$), crescendo a partir desse valor. Assim, o algoritmo de modelagem está indicando, para essa série, modelo auto-regressivo não-linear como o mais apropriado. A comparação do gráfico da curva $EQMN(k)$, para a série BRASFUND com os correspondentes gráficos para as séries autoregressivas não-lineares com ruído na figura 4.2.b, capítulo 4, reforça essa conclusão.

5.1.2 Dimensão do modelo

Os gráficos das curvas $EQMN(m)$, $m=1, \dots, 10$, para as séries BRASFUND, WHM e VALE, estão na figura 5.2. As curvas $EQMN(m)$ indicam dimensão de imersão m^* (dimensão do modelo) igual a 1, para as séries WHM e VALE, e dimensão de imersão m^* igual a 3, para a série BRASFUND. Para a série IBOVESPA, tem-se $EQMN(m) \approx 1.0$, $m=1, \dots, 10$. O menor valor e o maior valor do $EQMN(m)$ são, respectivamente, $EQMN(3)=1.0103$ e $EQMN(8)=1.047$. Adotou-se, então, para essa série a "dimensão de imersão $m=3$ ".

No quadro 5.1, estão os valores ótimos dos parâmetros m e k que minimizam o $EQMN(m, k)$. Observa-se que os erros são, em geral, elevados.

¹ Por exemplo: $T_2=300$, $k=460$ e $m=10$, implica em resolver 300 problemas de mínimos quadrados lineares com 10 variáveis explicativas, cada uma com comprimento 460.

QUADRO 5.1-VALORES ÓTIMOS DOS PARÂMETROS m^* E k^* QUE MINIMIZAM $EQMN(m,k)$

SÉRIE	Parâmetros		$EQMN(m^*,k^*)$
	m^*	k^*	
IBOVESPA	3	-	1.010
BRASFUND	3	220	0.735
WHM	1	460	0.778
VALE	1	460	0.727

m^* =dimensão de imersão, k^* =número de pontos em cada vizinhança(=número de pontos utilizados para obter cada uma das 300 previsões), $EQMN(m^*,k^*)$ =erro quadrático médio normalizado mínimo.

5.1.3 Testes estatísticos

O algoritmo de modelagem indica que as séries IBOVESPA, WHM e VALE, não são geradas por sistemas auto-regressivos não-lineares, isto é, não são processos não-lineares na média. Para a série BRASFUND, existe alguma evidência de não-linearidades na média. Modelos não-lineares na média e na variância são apresentados no apêndice A, Testes Estatísticos. Nesse apêndice estão, também, resumidos os testes estatísticos do terceiro momento incondicional e BDS.

Para investigar não-linearidades nessas séries, de forma alternativa à desenvolvida com o algoritmo de modelagem, são empregados os testes do terceiro momento incondicional(Hsieh, 1989), de McLeod-Li(McLeod and Li,1983) e BDS(Scheinkman,1990).

QUADRO 5.2-ESTATÍSTICAS BÁSICAS DAS SÉRIES FINANCEIRAS

SÉRIE	ESTATÍSTICAS					
	N.OBS.	MÉDIA	VAR.	DP.	ASS.	CURT.
IBOVESPA	1598	7.51E-3	2.12E-3	0.0461	0.6100	7.3598
BRASFUND	1579	3.88E-4	1.87E-3	0.0432	0.0141	2.6492
WHM	1307	7.90E-3	4.11E-3	0.0641	-0.2121	24.0518
VALE	1307	6.68E-3	5.92E-3	0.0769	0.2973	7.3154

CURT.=CURTOSE-3.0

FIGURA 5.1. Gráficos das curvas EQMN(k), $k=20, \dots, 460$, para as séries financeiras (retornos) BRASFUND, WHM e VALE obtidas através da regressão não-paramétrica LWR com os parâmetros, $p=1, \tau=1, T_2=300, m=1, \dots, 10, T_1$ e T_2 com os valores em 5.1. As curvas apresentadas são as curvas com o menor erro quadrático médio normalizado.

BRASFUND ———, WHM - - - - , VALE

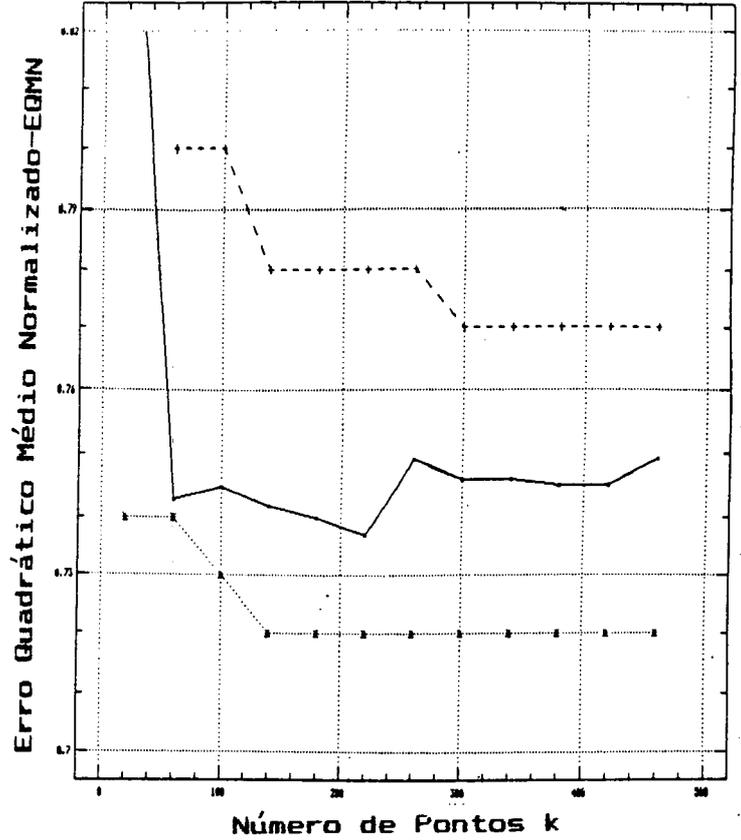
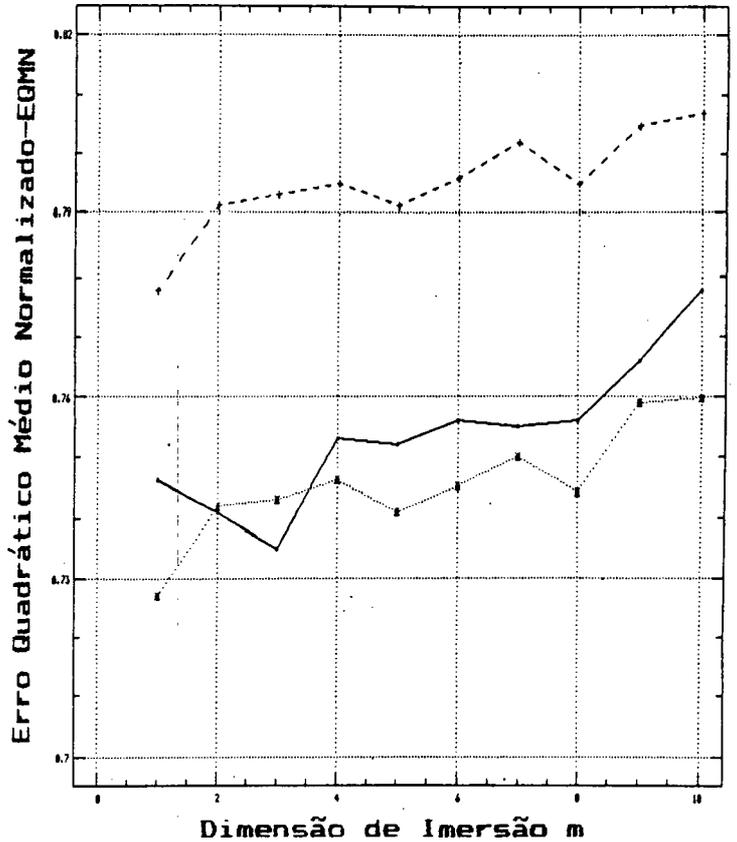


FIGURA 5.2. Gráficos das curvas EQMN(m), $m=1, \dots, 10$, para as séries financeiras (retornos) BRASFUND, WHM e VALE obtidas através da regressão não-paramétrica LWR com os parâmetros, $p=1, \tau=1, T_2=300, k=20, \dots, 460, T_1$ e T_2 com os valores dados na seção 5.1. As curvas apresentadas são as curvas com o menor erro quadrático médio normalizado.

BRASFUND ———, WHM - - - - , VALE



Teste do terceiro momento incondicional

O teste do terceiro momento incondicional(TM) pode apontar se as não-linearidades em séries temporais são na média ou na variância. O teste é aplicado aos resíduos do modelo linear ajustado à série. Com a metodologia de Box-Jenkins(Box and Jenkins,1976) foram identificados, estimados e checados os modelos lineares que melhor se ajustam a essas quatro séries financeiras. No quadro 5.4, estão as estatísticas de Portematau para os modelos lineares ajustados às séries. Os modelos lineares para as séries BRASFUND e WHM não têm boa especificação. As probabilidades de que os resíduos desses modelos lineares têm de serem independentes é muito baixa. Tem-se 29% para a série BRASFUND e 19% para a série WHM. Uma explicação possível para a baixa qualidade dos modelos lineares é a existência de auto-correlações de longo prazo nessas séries.

As estatísticas do teste do terceiro momento incondicional, aplicados aos resíduos do modelos lineares identificados para as quatro séries financeiras, estão no quadro 5.3. O teste não identifica não-linearidades na média, confirmando a conclusão obtida com o algoritmo de modelagem para as séries IBOVSPA, WHM e VALE. Para a série BRASFUND o teste do terceiro momento e o algoritmo de modelagem dão conclusões diferentes. O primeiro não apresenta estatísticas significativas, resultados no quadro 5.3, enquanto o segundo aponta evidências de variáveis auto-regressivas não-lineares nessa série.

Teste de McLeod-Li

O teste de McLeod-Li(TML) analisa as autocorrelações do processo dado pelos resíduos ao quadrado do modelo linear, para detectar não-linearidades na série temporal.

As estatísticas do teste TML, aplicado aos processos dos resíduos ao quadrado dos modelos lineares ajustados às séries financeiras, estão no quadro 5.4. As estatísticas significativas, a um nível de significância de 5%, mostram que as quatro séries são observadas de sistemas dinâmicos não-lineares.

Teste BDS

Rothman(1992,p.S189) argumenta que o teste estatístico BDS é um teste apropriado para séries financeiras, por não exigir a existência de momentos de ordem superior² para deduzir sua distribuição assintótica, e cita trabalhos recentes que apontam que momentos incondicionais de ordem elevada não existem para

séries financeiras típicas. Além disso, o teste BDS é muito sensível a desvios da distribuição iid. Por essas razões, as conclusões obtidas a partir desse teste são priorizadas nesta seção.

As estatísticas do teste BDS para $\epsilon=1.0$ DP e $m=2, \dots, 10$, aplicado aos resíduos dos modelos lineares ajustados às séries financeiras em análise, estão no quadro 5.5. As estatísticas significativas, valor crítico 2.57 a um nível de significância de 1%, indicam que essas séries são observadas de sistemas dinâmicos não-lineares.

5.1.4 Dimensão de correlação

A dimensão de correlação não é, propriamente, um teste estatístico. Mas, para processos iid., a dimensão de correlação ν segue a lei de escala $\nu(m)=m$, para todo m . Para os sistemas determinísticos caóticos com baixo nível de ruído, $\nu(m)$ torna-se independente de m , com m crescente.

Pretende-se, na discussão a seguir, comparar visualmente essa lei de escala com as dimensões de correlação dos resíduos dos modelos lineares ajustados às séries financeiras em análise. Na figura 5.3, estão os gráficos da lei de escala $\nu(m)=m$, $m=1, \dots, 10$, e os gráficos da dimensão de correlação dos resíduos dos modelos lineares, estimadas através do algoritmo (2.7), capítulo 2, com $\epsilon=0.50$ DP. Para os resíduos dos modelos lineares ajustados às séries IBOVESPA, WHM e VALE, as dimensões de correlação crescem linearmente com a dimensão de imersão, conforme mostram os resultados no quadro 5.6.

Para os resíduos do modelo linear ajustado à série BRASFUND tem-se $\nu(m) \approx 0.35$, $m=1, \dots, 10$, o que mostra para essa série, em termos da dimensão de correlação, características diferentes das outras três em análise. Essa série, como é indicado pelos algoritmos de modelagem e da dimensão de correlação, é um processo estocástico não-linear com baixa dimensão. Essa conclusão não deve surpreender. Ramsey et al. (1990, p.999) comentam que processos estocásticos com baixa dimensão são encontrados em muitas séries econômicas.

² O teste do terceiro momento incondicional, por exemplo, exige momentos iguais ou superiores a 4 (Hsieh, 1991, p.1861).

QUADRO 5.3-TESTE DO TERCEIRO MOMENTO INCONDICIONAL PARA OS RESÍDUOS DOS MODELOS LINEARES

DEF i,j	SÉRIE			
	IBOVESPA	BRASFUND	WHM	VALE
1,1	-0.62	1.84	-0.44	1.92
2,1	-0.45	-0.77	0.65	-0.74
2,2	0.18	1.47	1.30	-0.68
3,1	-2.51	1.80	-0.68	0.87
3,2	-0.57	-0.77	0.66	0.64
3,3	-0.75	-0.60	-0.70	-0.13
4,1	0.20	2.37	-0.34	0.76
4,2	-1.37	0.53	-0.33	-0.83
4,3	0.59	0.03	0.39	0.62
4,4	-1.48	-0.95	-0.66	-0.59
5,1	-1.18	0.23	-1.65	-0.17
5,2	-0.05	-0.48	-1.14	0.29
5,3	0.08	-0.30	0.02	-0.36
5,4	0.66	-0.86	-0.79	0.72
5,5	0.97	0.83	-0.74	-0.03

DEF=Defasagem. Valor crítico 2.57, a um nível de significância de 1%.

QUADRO 5.4-TESTE DE PORTEMATAU(TP) E DE MCLEOD-LI(TML) PARA AS SÉRIES ORIGINAIS E PARA OS RESÍDUOS DOS MODELOS LINEARES.

SÉRIE	DADOS BRUTOS		RESÍDUOS DO MODELO LINEAR			
	TP	TML	TP	Pr.	TML	Pr.
IBOVESPA	57.45	121.32	11.73	0.62	119.40	0.0
BRASFUND	112.53	138.53	13.01	0.29	96.25	0.0
WHM	30.30	52.75	17.22	0.18	38.61	0.0
VALE	161.97	121.11	7.58	0.67	53.82	0.0

Valores críticos, a um nível de significância de 5%, para o teste TP das 20 primeiras autocorrelações: IBOVESPA=25.00, BRASFUND=21.03, WHM=23.68 e VALE=19.68. Valor crítico para o teste TML, a um nível de significância de 5%, 31.41. Pr. refere-se a probabilidade dos resíduos serem ruído branco, considerando-se as autocorrelações remanescentes.

QUADRO 5.5-ESTATÍSTICA BDS(1.0 DP) PARA OS RESÍDUOS DOS MODELOS LINEARES

m	RESÍDUOS			
	IBOVESPA	BRASFUND	WHM	VALE
2	6.28	3.50	7.91	7.65
3	7.81	5.10	8.57	9.05
4	9.39	6.18	9.37	9.69
5	10.58	7.16	9.88	10.65
6	11.06	8.11	10.53	10.59
7	12.45	8.72	11.11	11.45
8	13.34	9.87	11.50	11.57
9	14.33	10.85	12.14	11.90
10	15.34	12.37	12.99	12.54

Valor crítico 2.57, a um nível de significância de 1%. O valor de m é a dimensão de imersão.

QUADRO 5.6-DECLIVIDADES DAS RETAS AJUSTADAS ÀS DIMENSÕES DE CORRELAÇÃO DOS RESÍDUOS DOS MODELOS LINEARES

RESÍDUOS	DECLIVIDADE	R	R ²
IBOVESPA	0.81(0.01)	0.99	0.99
WHM	0.66(0.01)	0.99	0.99
VALE	0.53(0.01)	0.99	0.99

O valor entre parênteses é o erro padrão da estimativa da declividade. R é o coeficiente de correlação e R² é o coeficiente de determinação da regressão linear.

5.2 O MODELO ARCH(p)

Os testes estatísticos de McLeod-Li e BDS permitem concluir que as quatro séries analisadas apresentam fortes evidências de não-linearidades. O algoritmo de modelagem indica que as não-linearidades nas séries IBOVESPA, WHM e VALE, não são devidas a variáveis auto-regressivas não-lineares, isto é, não são não-lineares na média, e que existe alguma evidência dessas variáveis no processo gerador da série BRASFUND. Uma alternativa aos modelos não-lineares na média é dada pelos modelos não-lineares na variância (Bollerslev, 1986, 1987; Nelson, 1991).

A seguir, procura-se investigar se as não-linearidades diagnosticadas nessas séries financeiras podem ser explicadas

através dos modelos de variância condicional ARCH(p), os quais pertencem ao grupo dos modelos não-lineares na variância.

5.2.1 Definição

Considere-se o modelo dinâmico de regressão linear

$$y_t = x_t' \beta + \varepsilon_t. \quad (5.1)$$

O erro $\{\varepsilon_t, t=1, \dots, T\}$ é gerado por um processo ARCH(p), segundo Engle (1982, p.988), se os resíduos padronizados

$$z_t = \varepsilon_t / h_t^{1/2} \quad (5.2)$$

têm distribuição $N(0,1)$, onde

$$h_t = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \dots + \alpha_p \varepsilon_{t-p}^2, \quad (5.3)$$

com

$$\alpha_0 > 0, \alpha_i \geq 0, i = 1, \dots, p, \quad (5.4)$$

para assegurar com probabilidade 1.0 que h_t é positiva.

5.2.2 Identificação e checagem dos modelos

O processo dos erros ao quadrado $\{\varepsilon_t^2, t=1, \dots, T\}$ pode ser utilizado para identificar e checar os processos ARCH(p). Bollerslev (1988, p.123) prova que esse processo tem a mesma estrutura de autocorrelações e correlações parciais de um AR(p). Decorre, da identificação dos modelos não-lineares ARCH(p) com os modelos lineares AR(p), que a metodologia Box-Jenkins pode ser utilizada nas etapas de identificação e checagem do modelo.

Um teste para a adequação do modelo é dado pela estatística de Portmanteau (Bollerslev, 1986, p.315).

$$Q_{EE} = T(T+2) \sum_{i=1}^N r_{EE}^2(i) / (T-i), \quad (5.5)$$

onde $r_{EE}(i)$ é a i -ésima autocorrelação do processo dado pelos resíduos padronizados ao quadrado

$$z_t^2 = \varepsilon_t^2 / h_t \quad (5.6)$$

obtidos de (5.2).

Quando os z_t são independentes, a estatística (5.5) tem distribuição assintótica ao qui-quadrado com N graus de liberdade.

5.2.3 Estimação dos modelos ARCH(p)

Embora os modelos ARCH(p) possam ser identificados com os modelos lineares AR(p), a estimação desses modelos não pode ser feita pela metodologia de Box-Jenkins, pois a equação da variância (5.3) com a condição (5.4) mostra que, estimar os coeficientes de um ARCH(p), é um problema de otimização com restrições de não-negatividade.

Os coeficientes em (5.3) podem ser estimados maximizando o logaritmo da função de verossimilhança, com as restrições de não-negatividade (5.4).

Para a implementação do processo de estimação dos coeficientes de um ARCH(p), neste trabalho, admite-se que $\{\varepsilon_t, t=1, \dots, T\}$ é observável, isto é, $\{\varepsilon_t\}$ foi estimado de forma consistente através de algum outro método de regressão (metodologia Box-Jenkins por exemplo). Engle (1982, p.990) deduz que a função log-máxima-verossimilhança para o processo definido em (5.2) e (5.3), a menos de alguma constante é:

$$L(\theta) = (1/T) \sum_{t=p+1}^T l_t(\theta), \quad (5.7)$$

com

$$l_t(\theta) = -1/2 \log h_t^2 - 1/2 \varepsilon_t^2 / h_t^2, \quad (5.8)$$

onde $\theta = (\alpha_0, \alpha_1, \dots, \alpha_p)$ é o vetor dos coeficientes a ser estimado maximizando $L(\theta)$ em (5.7), com as restrições de não-negatividade $\alpha_0 > 0$, $\alpha_i \geq 0$, $i=1, \dots, p$, dadas em (5.4).

5.3 MODELOS ARCH(p) ESTIMADOS

5.3.1 Coeficientes e erros padrões

Os coeficientes e os erros padrões dos modelos ARCH(p) estimados de (5.7), para os resíduos dos modelos lineares ajustados às séries financeiras em análise, estão no quadro 5.7. Os coeficientes são, em geral, significativos.

QUADRO 5.7—COEFICIENTES E ERROS PADRÕES DOS MODELOS ARCH(p) ESTIMADOS PARA AS SÉRIES FINANCEIRAS

COEFICIENTES	SÉRIE			
	IBOVESPA	BRASFUND	WHM	VALE
α_0	0.00087 (0.00008)	0.00142 (0.00016)	0.00351 (0.00025)	0.00387 (0.00048)
α_1	0.23777 (0.02326)	0.12057 (0.02170)	0.15498 (0.03140)	0.04851 (0.01985)
α_2	0.10353 (0.01996)	0.00194 (0.04300)		0.12743 (0.02474)
α_3	0.13366 (0.02653)	0.06909 (0.01752)		0.06346 (0.02215)
α_4	0.02189 (0.03218)			
α_5	0.10083 (0.02927)			

Os valores entre parênteses são os erros padrões das estimativas dos coeficientes.

As estatísticas básicas para os modelos ARCH estimados estão no quadro 5.8. A média e o desvio padrão são, respectivamente, muito próximos de 0.0 e de 1.0, como é exigido pela definição desses modelos.

QUADRO 5.8—ESTATÍSTICAS BÁSICAS DOS RESÍDUOS PADRONIZADOS DOS MODELOS ARCH(p) ESTIMADOS PARA AS SÉRIES FINANCEIRAS

ESTATÍSTICAS	SÉRIE			
	IBOVESPA	BRASFUND	WHM	VALE
Média	-0.00955	-0.00309	-0.0051	-0.00490
Variância	0.98898	0.99859	1.0024	1.00156
D. Padrão	0.99488	0.99929	1.0011	1.00077
Assimetria	0.17180	-0.11430	0.5044	0.06348
Curtose	1.08246	3.03272	27.4942	6.33083

5.3.2 Testes estatísticos

Teste de McLeod-Li

Os resultados do teste estatístico de McLeod-Li(TML),

aplicado ao processo dos resíduos padronizados ao quadrado (5.6), dos modelos ARCH(p) estimados para os resíduos dos modelos lineares ajustados às séries financeiras, estão no quadro 5.9.

QUADRO 5.9—TESTE DE MCLEOD AND LI(TML) PARA OS RESÍDUOS PADRONIZADOS DOS ARCH(p)

RESÍDUOS		TML	Pr.
IBOVESPA	ARCH(5)	17.70	54.0
BRASFUND	ARCH(3)	42.13	0.0
WHM	ARCH(1)	11.54	90.0
VALE	ARCH(3)	7.24	99.0

Valor crítico, a um nível de significância de 5%, para o teste TML das 20 primeiras autocorrelações, 31.41. Pr. refere-se à probabilidade dos resíduos serem ruído branco, considerando-se as autocorrelações remanescentes.

As estatísticas indicam que, para as séries IBOVESPA, WHM e VALE, as não-linearidades podem ser explicadas por esses modelos não-lineares na variância. Em particular, as probabilidades(Pr., quadro 5.9) dos processos dos resíduos padronizados ao quadrado, para as séries WHM(Pr.=90%) e VALE(Pr.=99%), de serem processos independentes são muito boas. Para a série BRASFUND, a probabilidade do processo dos resíduos padronizados ao quadrado ser independente é próxima de 0.0, mostrando que as não-linearidades nessa série não são explicadas por esse tipo de modelo, fornecendo mais uma evidência de que essa série é melhor modelada por um modelo estocástico auto-regressivo não-linear.

Embora os testes estatísticos TML, quadro 5.9, validem os modelos ARCH para as séries WHM e VALE, esses modelos falham nas estatísticas básicas, quadro 5.8. A curtose do processo dos resíduos padronizados é muito elevada para os dois modelos. Tem-se 27.49 e 6.33, respectivamente, para os resíduos padronizados da série WHM e VALE(compare-se com os correspondentes valores das séries originais, no quadro 5.2), mostrando que a distribuição desses resíduos não é $N(0,1)$, como é exigido na especificação dos modelos ARCH.

FIGURA 5.3. Dimensão de correlação, estimada através de (2.7 e 2.8, seção 2.2.3), para os resíduos dos modelos lineares ajustados às séries financeiras (retornos) IBOVESPA, BRASFUND, WMM e VALE; dimensão de correlação da lei de escala $\nu(m)=m$ (seção 2.1.4).

IBOVESPA-----, BRASFUND....., WMM-----
VALE-----, $\nu(m)=m$ ———

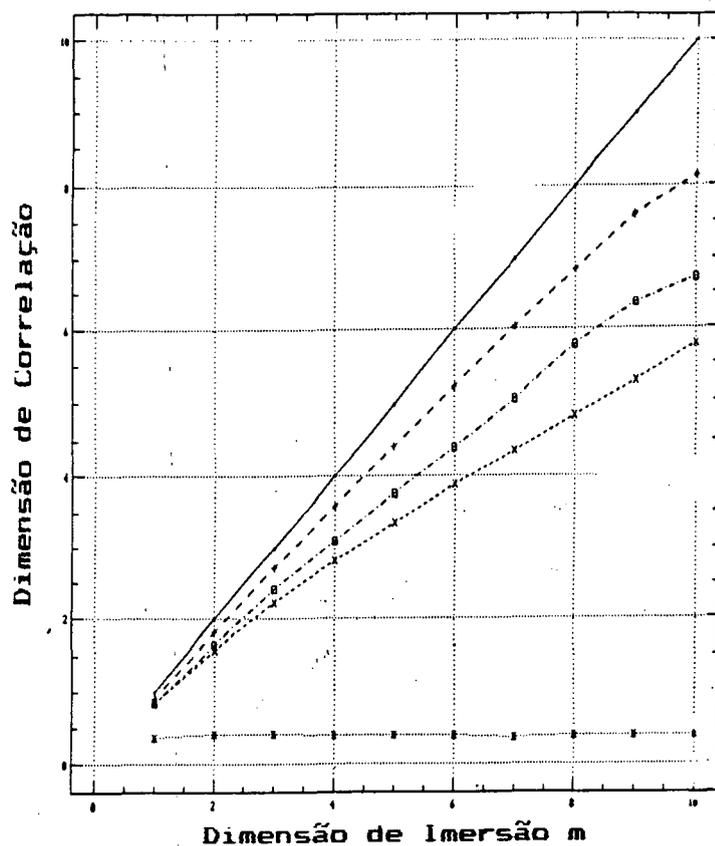
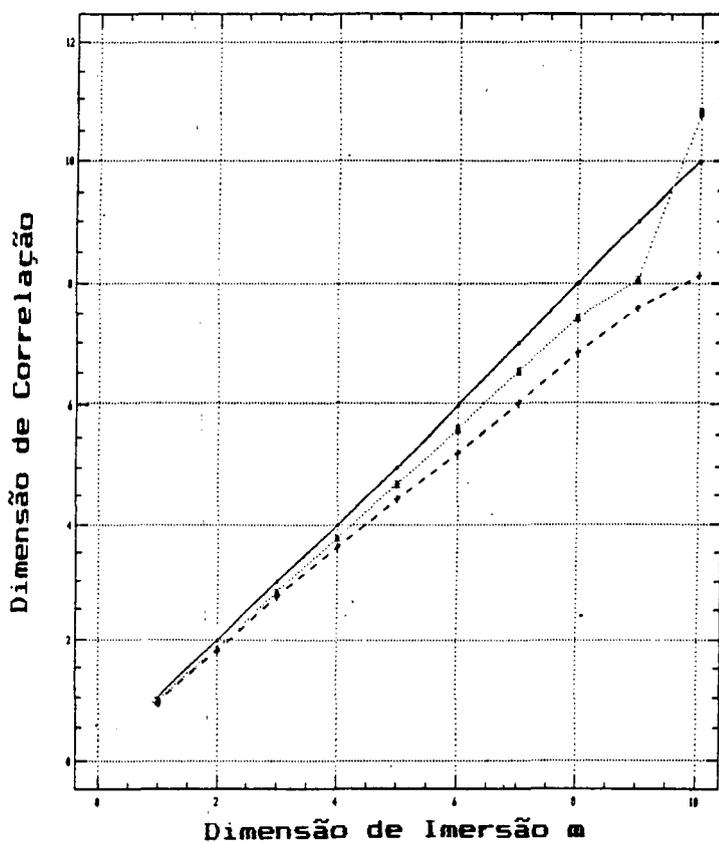


FIGURA 5.4. Dimensão de correlação, estimada através de (2.7 e 2.8, seção 2.2.3), para os resíduos dos modelos linear AR(5) e ARCH(5) ajustados à série (retornos) IBOVESPA; dimensão de correlação da lei de escala $\nu(m)=m$ (seção 2.1.4).

AR(5)-----, ARCH(5)....., $\nu(m)=m$ ———



Teste BDS

O teste BDS, para $m=2, \dots, 10$, e $\epsilon=1.0$ DP aplicado aos resíduos padronizados das quatro séries estão no quadro 5.10. Esse teste aponta que apenas os resíduos do ARCH(5) ajustado à série IBOVESPA são iid.

QUADRO 5.10—ESTATÍSTICA BDS(1.0 DP) PARA OS RESÍDUOS PADRONIZADOS DOS ARCH(p)

m	SÉRIE			
	IBOVESPA	BRASFUND	WHM	VALE
2	-1.21	1.93	2.85	4.29
3	-1.34	3.34	4.28	4.40
4	-1.37	4.65	5.54	4.47
5	-1.24	5.69	6.26	5.04
6	-1.41	6.60	7.00	5.39
7	-1.44	7.20	7.60	5.62
8	-1.41	8.38	7.96	5.57
10	-1.02	10.75	9.15	5.90

Valor crítico 2.57, a um nível de significância de 1%. O valor de m é a dimensão de imersão.

As conclusões diferentes dadas pelos testes BDS e TML podem ter origens no fato de que independentes e iid. não são sinônimos³, a não ser para processos gaussianos. Os resíduos dos modelos ARCH(p) ajustados às séries WHM e VALE, podem ser independentes e não distribuídos identicamente. O teste TML captura apenas independência. Outra possibilidade é que o teste BDS esteja apontando dependência linear nos resíduos. Observe-se, no quadro 5.4, que o modelo linear para WHM não é bom.

5.3 Dimensão de correlação

Na figura 5.4, são comparadas visualmente as dimensões de correlação dos resíduos do modelo linear AR(5) ajustado à série IBOVESPA, e dos resíduos padronizados do ARCH(5) ajustado a essa série, com a lei de escala $\nu=m$, que vale para processos iid. As

³ Ruído branco implica em processo independente, mas a recíproca não é verdadeira (Hinich and Patterson, 1985, p.70).

retas ajustadas às dimensões de correlação têm declividade: 0.81 para os resíduos do modelo linear e 0.99(0.91 abandonando o último ponto) para os resíduos padronizados do ARCH(5). Essa comparação, que não tem caráter estatístico, ilustra como o ARCH(5) estimado para a série IBOVESPA melhora a distribuição dos resíduos dessa série, aproximando-os da lei de escala $\nu(m)=m$.

CONCLUSÃO

A análise de sistemas dinâmicos não-lineares através da dimensão de correlação e do expoente de Lyapunov, como é desenvolvida tradicionalmente na literatura que investiga caos determinísticos em séries temporais do mundo real, exigem séries temporais com grande comprimento e baixo nível de ruído.

Quando aplicadas com sucesso, essas técnicas calculam invariantes geométricos e dinâmicos associados ao sistema não-linear, ignorando o problema fundamental em séries temporais com relação à previsão. Além disso, essas técnicas são aplicadas sem análise prévia da sua adequação a dada série temporal, podendo conduzir, às vezes, a resultados duvidosos (Grassberger, 1986), pois sistemas estocásticos lineares e sistemas estocásticos não-lineares podem ser identificados, erradamente, como sendo sistemas determinísticos caóticos.

As regressões não-paramétricas locais e globais permitem um ataque direto a esses problemas. O algoritmo de modelagem obtido através de aproximações locais (LWR) permite decidir qual das três modelagens, a linear auto-regressiva estocástica, a não-linear auto-regressiva estocástica ou a não-linear determinística, é a mais apropriada para dada série temporal. Para as séries identificadas como não-lineares determinísticas, é possível testar a presença de caos com as curvas EQMN(k) e EQMN(m). Para as séries identificadas como estocásticas auto-regressivas não-lineares e, também, para as determinísticas não-lineares, é possível analisar a sensibilidade às condições iniciais, estimando o expoente de Lyapunov dominante através de aproximações globais (rede neural). Esses testes e estimativas podem ser feitos com número de dados substancialmente menor que o exigido pelos algoritmos clássicos de análise de sistemas dinâmicos não-lineares.

Os resultados obtidos com a aplicação do algoritmo de modelagem, às séries financeiras brasileiras e à série das Manchas Solares, mostram a viabilidade de se analisar não-linearidades em séries temporais do mundo real com esse algoritmo. Além disso, a análise de séries temporais desenvolvida neste trabalho, com o algoritmo de modelagem, permite abandonar a dicotomia presente nos trabalhos que investigam caos em séries temporais do mundo real: o sistema é determinístico caótico ou é

ruído, através de uma aproximação do problema mais próxima da metodologia das Séries Temporais, que é a modelagem da série através de modelo estocástico auto-regressivo não-linear e a investigação da sensibilidade às condições iniciais dos modelos obtidos, a procura de evidências de caos determinístico quando o nível de ruído no modelo é muito pequeno.

A dimensão de imersão ideal, que o algoritmo de modelagem permite determinar, é importante não só no contexto desses modelos, permitindo minimizar os erros das previsões de curto prazo, mas também para os algoritmos clássicos de análise de sistemas dinâmicos não-lineares. O conhecimento da dimensão de imersão ideal maximiza as informações dadas pela série temporal, evitando trabalhar com dimensão de imersão crescente, procedimento que conduz à dimensão de imersão com alto valor, muitas vezes não suportada por séries com pequeno comprimento.

Embora invariantes geométricos e dinâmicos possam ser estimados para as séries determinísticas caóticas através dos modelos de previsão com aproximações locais, para essas séries o ponto importante desses modelos evidencia-se no diagnóstico de caos de baixa dimensão. Os modelos de previsão, que basicamente procuram reconstruir a dinâmica $f(F)$ do sistema dinâmico, são incapazes de diagnosticar, através das curvas EQMN(k) e EQMN(m), caos onde ele não exista, de forma diversa dos algoritmos de Grassberger-Procaccia (Grassberger, 1986) e de Wolf et al (Brock, 1986).

De forma objetiva, para aplicações práticas de sistemas dinâmicos com sensibilidade às condições iniciais, como é observado por Casdagli (1989) em relação aos sistemas determinísticos caóticos, é preferível um modelo de previsão monitorado pelo erro cometido, do que estimativas, quando é possível fazê-las, de invariantes dinâmicos e geométricos, que dependem fortemente do comprimento da série e do nível de ruído, e que são incapazes de fornecer informações quantitativas sobre a evolução do sistema dinâmico no tempo, a partir de dada condição inicial.

Além dos testes para diagnosticar caos determinístico e dos modelos de previsão para séries não-lineares auto-regressivas as técnicas apresentadas neste trabalho têm uma importante aplicação, na linha desenvolvida no capítulo 5: A análise da existência de variáveis auto-regressivas no sistema dinâmico gerador de séries observadas do mundo real. A aplicação dessa técnica não-linear mostra claramente que das séries financeiras analisadas (retornos) IBOVESPA, BRASFUND, WHM e VALE, somente a série BRASFUND apresenta evidências da existência de variáveis auto-regressivas não-lineares no sistema dinâmico gerador da série, enquanto as séries IBOVESPA, WHM e VALE são melhores modeladas, em termos de variáveis auto-regressivas, por modelos

auto-regressivos lineares.

A modelagem alternativa à auto-regressiva, através dos modelos de variância condicional (ARCH), permite concluir que esse tipo de modelo explica as não-linearidades diagnosticadas nas séries financeiras WHM e VALE por alguns dos testes estatísticos não-lineares utilizados neste trabalho. Em particular, o ARCH(5) ajustado à série IBOVESPA explica todas as não-linearidades capturadas pelos testes estatísticos não-lineares aqui utilizados. Por outro lado, os modelos ARCH não têm poder para explicar as não-linearidades observadas na série BRASFUND, fornecendo mais uma evidência de que essa série é melhor modelada por modelo auto-regressivo não-linear de dimensão 3, conforme é indicado pela análise da existência de variáveis auto-regressivas nessa série, desenvolvida através dos modelos não-lineares de previsão.

RECOMENDAÇÕES

Algumas direções que podem ser seguidas para aperfeiçoar as técnicas aqui apresentadas e, também, para testar a sua eficiência em outras aplicações são:

1) As aproximações lineares locais utilizadas nas regressões não-paramétricas podem ser substituídas por regressões de maior ordem, tal como a quadrática. A comparação com os resultados obtidos neste trabalho podem apontar qual a ordem da regressão não-paramétrica que possibilita previsões com menores erros quadráticos médios.

2) Os esquemas de pesos comumente utilizados na literatura de regressões não-paramétricas (tricúbica, por exemplo) não apresentaram, para os sistemas dinâmicos com sensibilidade às condições iniciais, desempenho superior ao esquema de pesos uniformes. Assim, novos esquemas de pesos para previsão de séries temporais observadas desses sistemas, através de técnicas de regressões não-paramétricas locais, podem ser pesquisados.

3) A decomposição do erro de previsão (erro quadrático médio normalizado) em duas partes: uma correspondente à sensibilidade às condições iniciais (Expoente de Lyapunov positivo) e a outra devido ao ruído presente na série temporal. E, a partir dessa decomposição, estabelecer critérios para decidir quando uma previsão é "boa" ou "ruim" (ver Apêndice C).

4) A conexão dos modelos de previsão para sistemas com sensibilidade às condições iniciais com novas áreas de pesquisa em sistemas dinâmicos com essas características (Por exemplo, Controle de Sistemas Caóticos) (ver, por exemplo, Ott et al., 1990; West, 1990; Glass et al., 1991).

5) A investigação da existência de variáveis auto-regressivas não-lineares em séries observadas de sistemas dinâmicos do mundo real e a modelagem dessas séries por modelos apresentados neste texto.

APÊNDICE A

TESTES ESTATÍSTICOS

Neste apêndice, são abordados testes estatísticos para identificar não-linearidades em séries temporais univariadas. O teste BDS, derivado da dimensão de correlação, permite testar a hipótese nula da série temporal ser um processo independente e identicamente distribuído, contra a hipótese alternativa da série ser gerada por processo não-linear. O teste do terceiro momento permite identificar se as não-linearidades são devidas a sistemas dinâmicos (processos) não-lineares na média ou na variância.

A.1 TESTE ESTATÍSTICO BDS

O teste BDS (Brock, Dechert e Scheinkman) (Hsieh, 1989; Scheinkman, 1990), desenvolvido a partir da integral de correlação, testa a hipótese nula de que os dados são independentes e identicamente distribuídos (iid.), tendo como alternativa processo estocástico linear, processo estocástico não-linear ou processo determinístico caótico.

O teste estatístico BDS é apresentado em Hsieh (1989, p.342), como segue:

Seja $\{x_t; t=1, \dots, T\}$ uma sequência de observações iid. Para dada dimensão de imersão m , construa-se as m -histórias (série de vetores m -dimensionais)

$$x_t^m = X_t = \{x_t, x_{t+1}, \dots, x_{t+m-1}\},$$

$t=1, \dots, T-m+1$, e estime-se a integral de correlação por (2.8), capítulo 2,

$$C(T, \epsilon, m) = \frac{2}{T_m(T_m-1)} \sum_{1 \leq t < s \leq T_m} I_{\epsilon}(X_t, X_s), \quad (A.1)$$

onde $T_m = T-m+1$ e I_{ϵ} é a função indicadora. Sob a hipótese nula de que $\{x_t\}$ é iid., com distribuição de probabilidade F , tem-se que:

$$C(T, \epsilon, m) \rightarrow C(T, \epsilon, 1)^m,$$

com probabilidade 1.0, quando $T \rightarrow \infty$. Além disso, com $T \rightarrow \infty$

$$T^{1/2}[C(T, \epsilon, m) - C(T, \epsilon, 1)^m] \sim N(0, \sigma^2(\epsilon, m)),$$

isto é, a distribuição limite é normal com média zero e variância

$$\sigma(\epsilon, m)^2 = 4[K(\epsilon)^m + 2 \sum_{j=1}^{m-1} K(\epsilon)^{m-1} C(\epsilon)^{2j} + (m-1)^2 C(\epsilon)^{2m} - m^2 K(\epsilon) C(\epsilon)^{2m-2}], \quad (A.2)$$

onde $C(\epsilon)$ e $K(\epsilon)$ podem ser estimados de forma consistente por $C(T, \epsilon, 1)$ em (A.1) e

$$K(T, \epsilon) =$$

$$= \frac{6}{T_m (T_m - 1) (T_m - 2)} \sum_{1 \leq t < s < r \leq T_m} I_\epsilon(X_t, X_s) I_\epsilon(X_s, X_r). \quad (A.3)$$

Portanto, a variância $\sigma(\epsilon, m)$ pode ser estimada de forma consistente pela fórmula (A.2) com $C(T, \epsilon, 1)$ e $K(T, \epsilon)$, respectivamente, no lugar de $C(\epsilon)$ e $K(\epsilon)$. Sob a hipótese nula, a estatística BDS

$$w(T, \epsilon, m) = T^{1/2}[C(T, \epsilon, m) - C(T, \epsilon, 1)^m] / \sigma(T, \epsilon, m)^2, \quad (A.4)$$

tem como distribuição limite a normal $N(0, 1)$.

A não aceitação da hipótese nula (valor crítico 2.57, a um nível de significância de 1%) indica algum tipo de dependência nos dados, que pode ser originada por processo estocástico linear, processo estocástico não-linear ou por processo não-linear determinístico. No caso da hipótese nula não ser aceita, análise complementar deve ser desenvolvida para decidir qual das três possibilidades alternativas está presente.

Brock and Baeck (1991, p.703) mostram que a estatística BDS pode ser utilizada para testar os resíduos de modelos da forma

$$x_t = F(x_{t-1}, x_{t-2}, \dots, x_{t-p}; \beta) + \varepsilon_t$$

com FEC^2 , os parâmetros $\beta = (\beta_1, \dots, \beta_2)$ estimados consistentemente e $\{\varepsilon_t\}$ iid. com média zero e variância unitária.

Assim, a estatística BDS pode ser utilizada para testar os resíduos de um $AR(p)$. Em particular, os resíduos de uma rede neural podem ser analisados através do teste BDS. Esse resultado permite, também, analisar não-linearidades em séries temporais, reduzindo a hipótese alternativa no teste BDS a processo não-linear estocástico ou determinístico, através da filtragem da série por modelo linear apropriado.

Uma propriedade importante do teste estatístico BDS é apontada por Rothman(1992,p.5189): O teste estatístico BDS não requer a existência de momentos de ordem superior para deduzir a sua distribuição assintótica. Os testes não-lineares bi-espectral de Hinich(1982,p.170) e o do terceiro momento incondicional (Hsieh,1991,p.1861), por exemplo, exigem a existência de momentos iguais ou superiores a 4.

A.2 TESTE DO TERCEIRO MOMENTO INCONDICIONAL

A.2.1 Modelos não-lineares

Hsieh(1989,p.354) separa os modelos estocásticos não-lineares em dois grandes grupos: modelos não-lineares na média e modelos não-lineares na variância, através da seguinte argumentação:

Seja $\{x_t\}$ os resíduos de um $AR(p)$ e considere os modelos da forma

$$x_t = f(x_{t-1}, x_{t-2}, \dots, x_{t-d}) + \varepsilon_t, \quad (A.5)$$

e os modelos da forma

$$x_t = f(x_{t-1}, x_{t-2}, \dots, x_{t-d}) \varepsilon_t, \quad (A.6)$$

onde $f: \mathbb{R}^d \rightarrow \mathbb{R}$ é não-linear e $\{\varepsilon_t\}$ é iid. com

$$E(\varepsilon_t | x_{t-1}, x_{t-2}, \dots, x_{t-d}) = 0,$$

isto é, $\{\varepsilon_t\}$ é independente dos valores passados de $\{x_t\}$ e

$$V(\varepsilon_t | x_{t-1}, x_{t-2}, \dots, x_{t-d}) = \sigma^2.$$

O primeiro modelo é dito com dependência aditiva, enquanto o segundo modelo é dito com dependência multiplicativa. Dependência aditiva significa que as não-linearidades se fazem presentes no sistema gerador da série temporal através da média (modelo não-linear na média), enquanto dependência multiplicativa significa que as não-linearidades se fazem presentes no sistema através da variância (modelos não-lineares na variância).

São exemplos do primeiro tipo os modelos: bilinear de Granger (BL) (Priestley, 1981, p.883) e auto-regressão por partes (TAR) (Tong and Lim, 1980). Os sistemas determinísticos caóticos e os sistemas caóticos com ruído (3.5), capítulo 3, também fazem parte desse grupo. Os modelos de variância heterocedástica tais como os ARCH (Engle, 1982) (5.2, capítulo 5), GARCH (Bollerslev, 1986) e EGARCH (Nelson, 1991), são exemplos de modelos não-lineares na variância.

A.2.2 O teste do terceiro momento incondicional

Para os modelos da forma (A.5),

$$E(x_t | x_{t-1}, x_{t-2}, \dots, x_{t-d}) = 0, \quad (\text{A.7})$$

enquanto para os modelos da forma (A.6),

$$E(x_t | x_{t-1}, x_{t-2}, \dots, x_{t-d}) = 0. \quad (\text{A.8})$$

As relações (A.7) e (A.8) permitem distinguir, como apontado por Hsieh (1989, 1991, p.1860), entre os dois tipos de modelos não-lineares (A.5) e (A.6), através do teste do terceiro momento incondicional, como segue:

A hipótese nula

$$E(f(x_{t-1}, x_{t-2}, \dots, x_{t-k})) = 0,$$

isto é, a existência de não-linearidade multiplicativa contra a hipótese alternativa

$$E(f(x_{t-1}, x_{t-2}, \dots, x_{t-k})) = 0,$$

da existência de não-linearidade aditiva, pode ser testada através do terceiro momento incondicional.

Seja $E(x_t, x_{t-i}, x_{t-j})$ o terceiro momento incondicional e $V(x_t) = \sigma^2$. Sob a hipótese nula,

$$\rho(i,j) = E(x_t, x_{t-i}, x_{t-j})/\sigma^3 = 0,$$

para todo $i,j>0$, contra a hipótese alternativa que para algum $i,j>0$

$$\rho(i,j) = 0.$$

Estimando-se $\rho(i,j)$ pelo momento amostral

$$r(i,j) = (1/T)[\sum x_t x_{t-i} x_{t-j}] / [(1/T)\sum x_t^2]^{3/2},$$

a distribuição assintótica da estatística $r(i,j)$ é dada por

$$N(0, w(i,j)/\sigma^6),$$

onde a variância pode ser estimada consistentemente por

$$\hat{w}(i,j)/\sigma^6 = (1/T)[\sum x_t^2 x_{t-i}^2 x_{t-j}^2] / [(1/T)\sum x_t^2]^3.$$

A estatística

$$T^{1/2} r(i,j) / [\hat{w}(i,j)/\sigma^6]^{1/2}$$

é, então, usada para testar quando o terceiro momento é diferente de zero. A rejeição da hipótese nula ocorre, a um nível de significância de 1%, se o valor absoluto dessa estatística é maior ou igual a 2.57.

Simulações em Hsieh(1989,p.358) mostram que o teste do terceiro momento tem bom poder para detectar não-linearidades na média e não-linearidades na variância.

APÊNDICE B

TEOREMA DE TAKENS

Neste apêndice, procura-se, através de exemplos, ilustrar o teorema de Takens(1981).

B.1 SISTEMA DINÂMICO DE DIMENSÃO 2

Considere-se o sistema de equações diferenciais de dimensão 2

$$dX/dt = \begin{cases} dx_1/dt = F_1(x_1, x_2) \\ dx_2/dt = F_2(x_1, x_2), \end{cases} \quad (B.1)$$

com $X=(x_1, x_2)' \in \mathbb{R}^2$ e $\partial F_1/\partial x_2=0$.

Schuster(1988, p.120) ilustra o teorema de Takens com o seguinte argumento:

Para $\tau=0$, os vetores

$$X(t_i) = X(t+i\tau) = [x_1(t+i\tau), x_2(t+i\tau)]',$$

i inteiro, são obtidos de

$$X(t) = [x_1(t), x_2(t)]'$$

univocamente. A correspondência(aplicação) entre os pontos $X(t_i)$ e $X(t)$ é injetora(1-a-1), pois as trajetórias do sistema (B.1) não se interceptam(unicidade do problema de valor inicial).

Considere-se a sequência de vetores

$$Y(t_i) = [y_1(t_i), y_2(t_i)]' = [x_1(t+i\tau), x_1(t+(i+1)\tau)]',$$

onde $t_i=t+i\tau$ e $i=1,2,\dots$

Tem-se,

$$y_1(t_i) = x_1(t+i\tau) \quad (B.2)$$

$$y_2(t_i) = x_1(t+(i+1)\tau) \int_{t+i\tau}^{t+(i+1)\tau} F_1(x_1(\alpha), x_2(\alpha)) d\alpha + x_1(t+i\tau) \approx \\ \tau F_1(x_1(t+i\tau), x_2(t+i\tau)) + x_1(t+i\tau) = \\ \tau F_1(x_1(t_i), x_2(t_i)) + x_1(t_i). \quad (B.3)$$

Considerando-se que $\partial F_1 / \partial x_2 = 0$, as igualdades (B.2) e (B.3) definem uma aplicação injetora entre $X(t_i)$ e $Y(t_i)$ (teorema da função inversa; isola-se em (B.3), $x_2(t_i)$ em função de $y_2(t_i)$). Assim, é razoável admitir-se que as seqüências $X(t_i)$ e $Y(t_i)$ tenham as "mesmas informações", tendo em vista que entre elas existe uma aplicação $S: \mathbb{R}^2 \rightarrow \mathbb{R}^2$ injetora.

B.2 CIRCUNFERÊNCIA UNITÁRIA

Como exemplo simples, considere-se o sistema de equações diferenciais:

$$\begin{bmatrix} dx_1/dt \\ dx_2/dt \end{bmatrix} = \begin{bmatrix} 0 & 2\pi \\ -2\pi & 0 \end{bmatrix} \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix} \quad (B.4)$$

com a condição inicial (0,1), cuja trajetória no plano de fase é a circunferência com centro em zero e raio unitário (circunferência unitária), pois $x_1(t) = \cos 2\pi t$ e $x_2(t) = \sin 2\pi t$, donde $[x_1(t)]^2 + [x_2(t)]^2 = 1$.

Tem-se, para $t_i = t + i\tau$ e $\tau = 1/4$

$$Y(t_i) = [x_1(t_i), x_1(t_i + \tau)]' = [x_1(t + i\tau), x_1(t + (i+1)\tau)]' = \\ [\sin 2\pi t_i, \sin(2\pi(t_i + 1/4))] ' = \\ [\sin 2\pi t_i, \cos 2\pi t_i] ' = X(t_i).$$

Para esse sistema, as seqüências $X(t_i)$ e $Y(t_i)$ são idênticas. A aplicação injetora dada por (B.2) e (B.3) é a aplicação identidade, mostrando que as seqüências são completamente equivalentes em termos de informação.

B.3 HÉNON

O teorema de Takens vale para sistemas de equações de diferenças (Brock, 1986, p.173). Considere-se, por exemplo, o sistema dinâmico de Hénon, com a notação $x_t = x(t)$ e $y_t = y(t)$,

$$\begin{cases} x(t) = 1 - a[x(t-1)]^2 + y(t-1) \\ y(t) = bx(t-1). \end{cases} \quad (\text{B.5})$$

Para $t_i = t + i\tau$ e $\tau = -1$, tem-se

$$Y(t_i) = [x(t_i), x(t_i + \tau)]' = [x(t_i), x(t_i - 1)]'$$

$$[x(t_i), y(t_i)/b]' = \begin{bmatrix} 1 & 0 \\ 0 & 1/b \end{bmatrix} [x_1(t_i), x_2(t_i)]'.$$

Assim,

$$Y(t_i) = \begin{bmatrix} 1 & 0 \\ 0 & 1/b \end{bmatrix} X(t_i),$$

é uma aplicação injetora que relaciona as seqüências $X(t_i)$ e $Y(t_i)$.

Pelo teorema de Takens (enunciado a seguir), as propriedades geométricas e dinâmicas dos sistemas (B.1), (B.4) e (B.5), com estados $X(t_i)$, que dependem de métricas, podem ser obtidas das correspondentes seqüências $Y(t_i)$, isto é, dos correspondentes sistemas de equações de diferenças construídos utilizando-se apenas uma das variáveis do sistema e defasagens dessa mesma variável. Esse procedimento é conhecido como "reconstrução do sistema dinâmico".

O resultado que Takens prova é, realmente, mais forte do que o resultado apresentado nos exemplos B.1, B.2 e B.3.

Teorema de Takens (Schuster, 1988, p.121):

Seja

$$dX(t)/dt = F(X(t))$$

um sistema dinâmico d -dimensional. Então,

$$Y(t) = [x_j(t), x_j(t+\tau), \dots, x_j(t+(2d+1)\tau)]',$$

onde $x_j(t)$ é uma componente arbitrária de $X(t) \in \mathbb{R}^d$, dá uma imersão suave (aplicação $f(X): U \subset \mathbb{R}^d \rightarrow \mathbb{R}^{2d+1}$, U aberto, com $Df(X): \mathbb{R}^d \rightarrow \mathbb{R}^{2d+1}$ injetora e contínua) do sistema d -dimensional em \mathbb{R}^d no espaço \mathbb{R}^{2d+1} , de tal forma que as propriedades métricas e dinâmicas do sistema no espaço \mathbb{R}^d , e as correspondentes propriedades métricas e dinâmicas da imersão do sistema no espaço \mathbb{R}^{2d+1} , são as mesmas.

Cronologicamente, Takens primeiro prova esse resultado para sistemas com tempo discreto (equações de diferenças, como está enunciado em 1.4.2.1, capítulo 1), isto é,

$$X(t) = F(X(t-1)) = F^t(X_0),$$

onde X_0 é a condição inicial e

$$F(X(t-1)) = F^t(X_0) = (F^{t-1} \circ F)(X_0)$$

para $t=1,2,\dots$,

estendendo esse resultado em 1981, para os sistemas de equações diferenciais.

APÊNDICE C

ENTROPIA MÉTRICA E EXPOENTE DE LYAPUNOV DOMINANTE

Neste apêndice, é investigada, para os sistemas do Grupo I, a lei de escala (4.7), que permite estimar o expoente de Lyapunov dominante ou a entropia métrica (soma dos expoentes de Lyapunov positivos), dependendo do tipo de regressão não-paramétrica utilizada para obter as aproximações locais. Como subproduto dessa investigação, dois problemas importantes para os modelos de previsão podem ser colocados.

C.1 REGRESSÃO NÃO-PARAMÉTRICA LWR

As curvas $EQMN(p)$ foram estimadas com $\tau=1$, $T=1500$, $T_1=1000$, $T_2=500$, m e k dados no quadro 4.2 e $p=1, \dots, 10$. As declividades das regressões (4.9) (entropia métrica) estão no quadro C.1. Para a Logística, Hénon e Lorenz, a entropia métrica é igual ao expoente de Lyapunov dominante (Abarbanel et al, 1990, p. 1791, 1792). Observe-se a concordância dos valores estimados de (4.8) no quadro C.1, para esses três sistemas dinâmicos, com os correspondentes valores estimados pelo algoritmo Global com redes neurais (quadro 4.1, capítulo 4).

Casdagli (1989, p.349) argumenta que:

A lei de escala (4.7) tende a superestimar os erros de previsão, provavelmente devido à taxa de expansão local no atrator, sendo necessário descartar previsões "ruins" para observá-la com fidelidade.

Segundo o autor, a taxa de expansão local é superenfaticada por duas razões:

- 1) a utilização do erro quadrático médio para estimar o erro de previsão, o qual é muito sensível a previsões ruins(?);
- 2) a tendência de existirem poucos pontos nas regiões de grande expansão do atrator(?), o que impossibilita previsões de boa qualidade nessas regiões,

e conjectura que a lei de escala dada por

$$MG(p) \approx Ce^{h(\alpha+1)p}, \quad (C.1)$$

onde $MG(p)$ é a média geométrica dos erros de previsão $e_t(T,m,p,k)$ em (4.3), e os valores de h e α os mesmos que em (4.7), é menos sensível a previsões ruins que a lei de escala (4.7), que utiliza o erro quadrático médio e, portanto, dá resultados mais próximos aos esperados.

De (C.1), tem-se:

$$\log MG(p) \approx \log C + h(\alpha+1)p, \quad (C.2)$$

o que permite estimar a entropia métrica ou o expoente de Lyapunov dominante, dependendo do tipo de regressão empregada no modelo de previsão.

A conjectura de Casdagli de que as estimativas do expoente de Lyapunov e da entropia métrica obtidas de (4.8) tendem a superestimar as estimativas obtidas através de (C.2), não é confirmada pelos resultados obtidos para os sistemas determinísticos caóticos neste texto (quadro C.1), obtidos sem descartar previsões (embora o programa computacional permita fazê-lo), os quais apontam em direção contrária para os sistemas Logística, Hénon, Mackey-Glass e Ikeda. Para o modelo de Lorenz, os valores estimados são próximos.

Existe, ainda, uma terceira possibilidade para explicar a origem das previsões ruins: para os modelos de previsão com aproximações locais lineares, é necessário algoritmo estável de mínimos quadrados, tal como decomposição em valores singulares. Esse algoritmo pode, eventualmente, apresentar problemas numéricos para séries determinísticas caóticas (a regressão linear é obtida de pontos que estão muito próximos). Como consequência, algumas previsões podem ser muito ruins, sendo necessário descartá-las.

A figura 4.5 apresenta os gráficos das curvas $EQMN(T,m,p,k) = EQMN(T)$, com $T_1=400,600,\dots,1200$, $T_2=300$, $p=1$, $\tau=1$, $m=3$ (Hénon), $m=4$ (Lorenz) e $k=2(m+1)$. Para os sistemas determinísticos caóticos com baixo nível de ruído, vale (4.9), donde

$$\log EQMN(T) \approx \log C + (\alpha/D) \log(k/T_1). \quad (C.3)$$

Assim, as curvas na figura 4.5 deveriam ser aproximadamente lineares. Observe-se que, na curva correspondente a Hénon, o

afastamento dessa lei de escala deve-se a "previsões ruins".

Introduz-se, assim, um fator subjetivo na utilização desses modelos de previsão para séries determinísticas caóticas, quando investigam-se as leis de escala (4.7) e (4.9): qual o critério usado (em Casdagli, por exemplo) para que uma previsão seja considerada como muito ruim? Quantas previsões serão descartadas? Qual é a origem da previsão ruim? É decorrente de problemas numéricos ou é consequência de flutuações bruscas intrínsecas à dinâmica caótica?

Essa discussão deixa em aberto dois importantes problemas, não disjuntos, para os modelos de previsão com aproximações lineares locais:

1) decompor o erro de previsão em duas partes¹: uma correspondente ao ruído e outra correspondente ao expoente de Lyapunov local (Wolf, 1990);

2) estabelecer critérios para decidir se uma previsão é "boa ou ruim" em função de 1).

QUADRO-C.1 ENTROPIA MÉTRICA h. REGRESSÃO NÃO-PARAMÉTRICA LWR

SISTEMA	ENTROPIA MÉTRICA h	
	EQMN	MG
Logística	0.6820 (0.0070, 0.99, 0.99)	0.7338 (0.0070, 0.99, 0.99)
Hénon	0.4108 (0.0264, 0.98, 0.96)	0.5003 (0.0229, 0.99, 0.98)
M.-Glass	0.1795 (0.0163, 0.96, 0.93)	0.1901 (0.0094, 0.99, 0.98)
Ikeda	0.3633 (0.0436, 0.94, 0.89)	0.4370 (0.0177, 0.99, 0.98)
Lorenz	0.2256 (0.0290, 0.93, 0.88)	0.2196 (0.0207, 0.96, 0.93)

Os valores entre parênteses são, respectivamente, o erro padrão da declividade (entropia métrica), o coeficiente de correlação R e o coeficiente de determinação R^2 das estimativas dadas pela regressão linear (4.9) e (C.2).

C.2 REGRESSÃO NÃO-PARAMÉTRICA NN

A análise anterior foi repetida integralmente para os mesmos sistemas, com a regressão não-paramétrica NN. Os resultados obtidos com essa regressão mais simples (quadro C.2) são próximos aos obtidos com a regressão LWR (quadro C.1), para a lei de escala (4.7). Para a lei de escala (C.1), os resultados obtidos (quadro C.2) são próximos aos estimados com o algoritmo GLOBAL (quadro 4.1, capítulo 4). Observe-se também, que os coeficientes de correlação (R) e determinação (R^2) de (4.8) e (C.2), obtidos com a regressão não-paramétrica NN, resultados no quadro C.2, são iguais ou superiores aos correspondentes coeficientes no quadro C.1, obtidos com a regressão não-paramétrica LWR. Em particular, deve-se destacar a concordância das estimativas do expoente de Lyapunov no quadro C.2, através da lei de escala (C.1), com as estimativas dadas pelo algoritmo WSSV², que é considerado um algoritmo muito bom para séries determinísticas caóticas com comprimento moderado e baixo nível de ruído.

¹ Um trabalho que aponta nessa direção é o de Yao and Tong (1994), o qual mostra que o erro quadrático médio pode ser decomposto em duas partes: uma devida à variância condicional e a outra resultante de pequenas perturbações nas condições iniciais.

² O algoritmo de Wolf et al. (WSSV) e a regressão não-paramétrica NN estão intimamente relacionados. A forma de selecionar pontos (para dada condição inicial), e obter os valores das trajetórias (previsões) a partir dos pontos selecionados são iguais no algoritmo WSSV e na regressão não-paramétrica NN, com esquema de pesos uniformes.

QUADRO C.2-EXPOENTE DE LYAPUNOV DOMINANTE. REGRESSÃO NÃO-PARAMÉTRICA NN

SISTEMA	EXPOENTE DE LYAPUNOV	
	EQMN	MG
Logística	0.6747 (0.0079, 0.99, 0.99)	0.6998 (0.0036, 0.99, 0.99)
Hénon	0.4037 (0.0135, 0.99, 0.99)	0.3994 (0.0038, 0.99, 0.99)
M.-Glass	0.1511 (0.0131, 0.97, 0.94)	0.1184 (0.0054, 0.99, 0.98)
Ikeda	0.3719 (0.0122, 0.99, 0.99)	0.2898 (0.0017, 0.99, 0.99)
Lorenz	0.2670 (0.0307, 0.95, 0.90)	0.1854 (0.0160, 0.97, 0.94)

Os valores entre parênteses são, respectivamente, o erro padrão da declividade(expoente de Lyapunov), o coeficiente de correlação R e o coeficiente de determinação R^2 das estimativas dadas pela regressão linear (4.9) e (C.2).

REFERÊNCIAS BIBLIOGRÁFICAS

- Abarbanel, H.D.I., Brown, R. and Kadtke, J.B., 1989, Prediction and system identification in chaotic nonlinear systems: time series with broadband spectra, *Physics Letters A*, 138, 401-408.
- , ——— and ———, 1990, Prediction in chaotic nonlinear systems: methods for time series with broadband fourier spectra, *Physical Review A*, 41, 1782-1807.
- Baumol, W.J. and Benhabib, J., 1989, Chaos: significance, mechanism and economic applications, *Journal of Economic Perspectives*, 3, 77-105.
- Ben-Mizrachi, A., Procaccia, I. and Grassberger, P., 1984, Characterization of experimental (noisy) strange attractors, *Physical Review A*, 29, 975-977.
- Bollerslev, T., 1986, Generalized autoregressive conditional heteroskedasticity, *Journal of Econometrics*, 31, 307-327.
- , 1987, A conditionally heteroskedastic time series model for speculative prices and rates of return, *The Review of Economics and Statistics*, 69, 542-547.
- , 1988, On the correlation structure for the generalized autoregressive conditional heteroskedastic process, *Journal of Time Series Analysis*, 9, 121-131.
- Box, G.E.P. and Jenkins, G.M., 1976, *Time series analysis, forecasting and control*, Holden-Day.
- Brock, W.A., 1986, Distinguishing random and deterministic systems: abridged version, *Journal of Economic Theory*, 40, 168-195.
- and Sayers, C., 1988, Is the business cycle characterized by deterministic chaos?, *Journal of Monetary Economics*, 22, 71-90.
- and Baek, E.G., 1991, Some theory of statistical inference for nonlinear science, *Review of Economic Studies*, 58, 697-716.
- Broomhead, D.S. and Jones, R., 1989, Time-series analysis, *Proc. R. Soc. Lond. A*, 423, 103-121.

- Cao, C.Q. and Tsay, R.S., 1992, Nonlinear time-series analysis of stock volatilities, *Journal of Applied Econometrics*, 7, S165-S185.
- Casdagli, M., 1989, Nonlinear prediction of chaotic time series, *Physica D*, 35, 335-356.
- , 1992, Chaos and deterministic versus stochastic non-linear modelling, *Journal of the Royal Statistical Society B*, 54, 303-328.
- Cheng, L. and Tong, H., 1992, On consistent nonparametric order determination and chaos, *Journal of the Royal Statistical Society B*, 42, 427-449.
- Cleveland, W.S., 1979, Robust locally weighted regression and smoothing scatterplots, *Journal of the American Statistical Association*, 74, 829-836.
- and Devlin, S.J., 1988, Locally weighted regression: an approach to regression analysis by local fitting, *Journal of the American Statistical Association*, 83, 596-610.
- , ——— and Grosse, E., 1988, Regression by local fitting, *Journal of Econometrics*, 37, 87-114.
- Dechert, W.D. and Gencay, R., 1992, Lyapunov exponents as a nonparametric diagnostic for stability analysis, *Journal of Applied Econometrics*, 7, S41-S60.
- Diebold, F.X. and Nason, J.A., 1990, Nonparametric exchange rate prediction, *Journal of International Economics*, 28, 315-332.
- Eckmann, J.P. and Ruelle, D., 1985, Ergodic theory of chaos and strange attractors, *Reviews of Modern Physics*, 57, 617-656.
- , Kamphorst, S.O., ——— and Ciliberto, S., 1986, Lyapunov exponents from time series, *Physical Review A*, 34, 4971-4979.
- Engle, R.F., 1982, Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation, *Econometrica*, 50, 987-1007.
- Farmer, J.D., 1982, Information dimension and the probabilistic structure of chaos, *Z. Naturforsch*, 37a, 1304-1325.
- , Ott, E. and Yorke, J.A., 1983, The dimension of chaotic attractors, *Physica D*, 7, 153-180.
- and Sidorowich, J.J., 1987, Predicting chaotic time series, *Physical Review Letters*, 59, 845-848.
- Frank, M.Z. and Stengos, T., 1988a, Chaotic dynamics in economic time-series, *Journal of Economic Surveys*, 2, 103-133.
- and ———, 1988b, Some evidence concerning macroeconomic chaos, *Journal of Monetary Economics*, 22, 423-428.

- Glass, L., Hunter, P. and McCulloch, A., 1991, *Theory of heart, Biomechanics, biophysics and nonlinear dynamics of cardiac function*, Springer-Verlang.
- Granger, C.W.J. and Newbold, P., 1976, Forecasting transformed series, *Journal of the Royal Statistical Society B*, 38, 189-203.
- Grassberger, P. and Procaccia, I., 1983a, Characterization of strange attractors, *Physical Review Letters*, 50, 346-349.
- and ———, 1983b, Measuring the strangeness of strange attractors, *Physica D*, 9, 189-208.
- , 1986, Do climatic attractors exist?, *Nature*, 323, 609-612.
- , 1990, An optimized box-assisted algorithm for fractal dimensions, *Physics Letters A*, 148, 63-68.
- Grebogi, C., Ott, E., Pelikan, S. and York, J.A., 1984, Strange attractors that are not chaotic, *Physica D*, 13, 261-268.
- Groot, G. and Wurtz, D., 1991a, Forecasting time series with connectionist nets: applications in statistics, signal processing and economics, *IPS Research Report*, no. 91-16.
- and ———, 1991b, Analysis of univariate time series with connectionist nets: a case study of two classical examples, *Neurocomputing*, 3, 177-192.
- Hénon, M., 1976, A two-dimensional mapping with a strange attractors, *Communication in Mathematical Physics*, 50, 69-77.
- Hsieh, D.A., 1988, The statistical properties of daily foreign exchange rates: 1974-1983, *Journal of International Economics*, 24, 129-145.
- , 1989, Testing for nonlinear dependence in daily foreign exchange rates, *Journal of Business*, 62, 339-368.
- , 1991, Chaos and nonlinear dynamics: application to financial markets, *Journal of Finance*, 46, 1839-1877.
- Hinich, M.J., 1982, Testing for gaussianity and linearity of a stationary time series, *Journal of Time Series Analysis*, 3, 169-176.
- , M.J. and Patterson, D.M., 1985, Evidence of nonlinearity in daily stock returns, *Journal of Business & Economic Statistics*, 3, 69-77.
- Koçak, H., 1986, *Differential and difference equation through computer experimentes*, Springer-Verlang.
- Lauwerier, H.A., 1986, *One-Dimensional Iterative Maps*. In *Chaos*, ed. A.V. Holden, Princeton University Press, Princeton, New Jersey.

- Liu, T., Granger, C.W. and Heller, W.P., 1992, Using the correlation exponent to decide whether an economic series is chaotic, *Journal of Applied Econometrics*, 7, S25-S39.
- Lorenz, N., 1963, Deterministic nonperiodic flow, *Journal of the Sciences*, 20, 130-141.
- Mandelbrot, B.B., 1982, *The fractal geometry of nature*, Freeman, San Francisco.
- McCaffrey, D.F., Ellner, S., Gallant, A.R. and Nychka, D.W., 1992, Estimating the Lyapunov exponent of a chaotic system with nonparametric regression, *Journal of the American Statistical Association*, 87, 682-695.
- Mackey, M.C. and Glass, L., 1977, Oscillation and chaos in physiological control systems, *Science*, 197, 287-289.
- May, R.M., 1976, Simple mathematical models with very complicated dynamics, *Nature*, 261, 459-467.
- McLeod, A.I. and Li, W.K., 1983, Diagnostic checking arma time series models using squared-residual autocorrelations, *Journal of Time Series Analysis*, 4, 269-273.
- Mizrach, B., 1992, Multivariate nearest-neighbour forecasts of EMS exchange rates, *Journal of Applied Econometrics*, 7, S151-S163.
- Nelson, D.B., 1991, Conditional heteroskedasticity in asset returns: a new approach, *Econometrica*, 59, 347-370.
- Nychka, D., Ellner, S., Gallant, A.R. and McCaffrey, D., 1992, Finding chaos in noisy systems, *Journal of the Royal Statistical Society B*, 399-426.
- Ott, E., Grebogi, C. and Yorke, J.A., 1990, Controlling chaos, *Physical Review Letters*, 64, 1196-1199.
- Packard, N.H., Crutchfield, J.P., Farmer, J.D. and Shaw, R.S., 1980, Geometry from a time series, *Physical Review Letters*, 45, 712-716.
- Press, W.H., Flannery, B.P., Teukolsky, S.A. and Vetterling, W.T., 1986, *Numerical Recipes, The art of scientific computing*, Cambridge University Press, Cambridge.
- Priestley, M.B. and Chao, M.T., 1972, Non-parametric function fitting, *Journal of the Royal Statistical Society B*, 34, 385-393.
- , 1980, State-dependent models: a general approach to non-linear time series analysis, *Journal of Time Series Analysis*, 1, 47-71.
- , 1981, *Spectral analysis and time series*, vol. 2, Academic Press.

- Ramsey, J.B. and Yuan, H., 1989, Bias and error bars in dimension calculation and their evaluation in some simple models, *Physics Letters A*, 134, 287-297.
- , Sayers, C.L. and Rothman, P., 1990, The statistical properties of dimension calculations using small data sets: some economic applications, *International Economic Review*, 31, 991-1020.
- Robinson, P.M., 1983, Nonparametric estimators for time series models, *Journal of Time Series Analysis*, 4, 185-208.
- Romeiras, F.J., Grebogi, C., Ott, E. and Dayawansa, W.P., 1992, Controlling chaotic dynamical systems, *Physica D*, 58, 165-192.
- Rothman, P., 1992, The comparative power of the TR test against simple threshold models, *Journal of Applied Econometrics*, 7, S187-S195.
- Ruelle, D., 1990, Deterministic chaos: the science and the fiction, *Proc. R. Soc. Lond. A*, 427, 241-248.
- Sano, M. and Sawada, Y., 1985, Measurement of the Lyapunov spectrum from a chaotic time series, *Physical Review Letters*, 55, 1082-1085.
- Scheinkman, J.A. and LeBaron, B., 1989, Nonlinear dynamics and stock returns, *Journal of Business*, 62, 311-337.
- , 1990, Nonlinearities in economic dynamics, *The Economic Journal*, 100, 33-48.
- Schuster, H.G., 1988, *Deterministic chaos, an introduction*, Cambridge.
- Smith, L.A., 1988, Intrinsic limits on dimension calculations, *Physics Letters A*, 133, 283-288.
- Smith, R.L., 1992, Estimating dimension in noisy chaotic time series, *Journal of the Royal Statistical Society B*, 54, 329-351.
- Stone, C.J., 1977, Consistent non-parametric regression, *Annals of Statistics*, 5, 595-645.
- Sugihara, G., Grenfell, B. and May, R.M., 1990, Distinguishing error from chaos in ecological time series, *Phil. Trans. R. Soc. Lond. B*, 330, 235-251.
- and May, R.M., 1990, Nonlinear forecasting as a way of distinguishing chaos from measurement error in time series, *Nature*, 344, 734-741.
- Theiler, J., 1986, Spurious dimension from correlation algorithms applied to limited time series data, *Physical Review A*, 34, 2427-2432.

- , 1987, Efficient algorithm for estimating the correlation dimension from a set of discrete points, *Physical Review A*, 36, 4456-4462.
- , 1990, Statistical precision of dimension estimators, *Physical Review A*, 41, 3038-3051.
- Tong, H. and Lim, K., 1980, Threshold autoregression, limit cycles and cyclical data, *Journal of the Royal Statistical Society B*, 42, 245-292.
- , 1990, *Non-linear time series, a dynamical system approach*, Clarendon Press Oxford, Oxford.
- West, B.J., 1990, *Fractal physiology and chaos in medicine*, World Scientific.
- White, H., 1989, Some asymptotic results for learning in simple hidden-layer feedforward network models, *Journal of the American Statistical Association*, 408, 1003-1013.
- Wolf, A.B., Swift, J.B., Swinney, H.L. and Vastano, J.A., 1985, Determining Lyapunov exponents from a time series, *Physica D*, 16, 285-317.
- Wolff, R.L.C., 1992, Local Lyapunov exponents: looking closely at chaos, *Journal of the Royal Statistical Society B*, 54, 353-371.
- Yakowitz, S.J., 1987, Nearest neighbor methods for time series analysis, *Journal of Time Series Analysis*, 8, 235-247.
- Yao, Q. and Tong, H., 1994, Quantifying the influence of initial values on non-linear prediction, *Journal of the Royal Statistical Society B*, 56, 701-725.