

UNIVERSIDADE DE ÉVORA

DEPARTAMENTO DE MATEMÁTICA



Análise não linear de séries temporais e aplicações à Psicologia

Elsa Rosa Arantes Lopes

MESTRADO EM MATEMÁTICA E APLICAÇÕES

Dissertação orientada pela Professora Doutora Sara Luísa Dimas Fernandes
e co-orientado pelo Professor Doutor Luís Manuel Ferreira da Silva

Dezembro de 2009

UNIVERSIDADE DE ÉVORA

DEPARTAMENTO DE MATEMÁTICA



Análise não linear de séries temporais e aplicações à Psicologia

Elsa Rosa Arantes Lopes

MESTRADO EM MATEMÁTICA E APLICAÇÕES



185045

Dissertação orientada pela Professora Doutora Sara Luísa Dimas Fernandes
e co-orientado pelo Professor Doutor Luís Manuel Ferreira da Silva

Dezembro de 2009

Dedico este trabalho aos meus pais.

Agradecimentos

Um dos melhores momentos no processo de escrever uma tese é aquele em que o autor tem a oportunidade de agradecer àqueles que o ajudaram, porque raramente um investigador faz o seu trabalho sozinho!

Os meus mais sinceros agradecimentos:

À Professora Doutora Sara Fernandes minha orientadora, que soube estimular, aconselhar, criticar e orientar de forma segura este trabalho até à sua conclusão, visando sempre o meu crescimento profissional e pessoal. Agradeço-lhe ainda todo o espaço que me disponibilizou e ainda todas as horas “gastas” para me orientar no caminho certo.

Ao Professor Doutor Luís Manuel Ferreira da Silva, co-orientador desta tese, pela sua disponibilidade, motivação, apoio, paciência, ajuda prestada e sugestões apresentadas.

Aos meus colegas e professores que fizeram do primeiro ano, um ano inesquecível e muito especial.

À minha amiga de mestrado, Daniela Brandão que nunca esquecerei pelos momentos que passamos juntas e as várias formas com que me ajudou durante a realização deste trabalho.

A todos os amigos que fiz e à minha família que sempre me apoiaram e encorajaram na realização deste trabalho.

Aos meus pais, que me ensinaram a lutar com dignidade, acreditaram no meu sonho e sempre desejaram o melhor para mim, pelo meu sucesso profissional e pela minha felicidade.

Muito especialmente ao meu marido, Alexandre, pelo apoio, paciência e ajuda infinita durante todo este trabalho. Sempre que necessário soube aconselhar e soube criticar, como sempre e em tudo na vida. Pelas alegrias, momentos felizes, desânimos, angústias e essencialmente pela compreensão.

“Nada se constrói sem que alguém tenha sonhado com isso, alguém tenha acreditado que isso fosse possível e alguém tenha querido que isso acontecesse!...”

Charles Kettering.

Análise não linear de séries temporais e aplicações à Psicologia

RESUMO:

Nesta dissertação estudamos as séries temporais que representam a complexa dinâmica do comportamento. Damos especial atenção às técnicas de dinâmica não linear. As técnicas fornecem-nos uma quantidade de índices quantitativos que servem para descrever as propriedades dinâmicas do sistema. Estes índices têm sido intensivamente usados nos últimos anos em aplicações práticas em Psicologia. Estudamos alguns conceitos básicos de dinâmica não linear, as características dos sistemas caóticos e algumas grandezas que caracterizam os sistemas dinâmicos, que incluem a dimensão fractal, que indica a complexidade de informação contida na série temporal, os expoentes de Lyapunov, que indicam a taxa com que pontos arbitrariamente próximos no espaço de fases da representação do espaço dinâmico, divergem ao longo do tempo, ou a entropia aproximada, que mede o grau de imprevisibilidade de uma série temporal. Esta informação pode então ser usada para compreender, e possivelmente prever, o comportamento.

PALAVRAS-CHAVE: Séries temporais, atractores, expoentes de Lyapunov, entropia, reconstrução do espaço de fases e dimensão de correlação.

Nonlinear analysis of time series and applications to Psychology

ABSTRACT:

In this thesis we studied the time series that represent the complex dynamic behavior. We focused on techniques of nonlinear dynamics. The techniques provide us a number of quantitative indices used to describe the dynamic properties of the system. These indices have been extensively used in recent years in practical applications in psychology. We studied some basic concepts of nonlinear dynamics, the characteristics of chaotic systems and some quantities that characterize the dynamic systems, including fractal dimension, indicating the complexity of information in the series, the Lyapunov exponents, which indicate the rate at that arbitrarily close points in phase space representation of a dynamic, vary over time, or the approximate entropy, which measures the degree of unpredictability of a series. This information can then be used to understand and possibly predict the behavior.

KEY-WORDS: Time Series, attractors, Lyapunov exponents, entropy, reconstruction of phase space and correlation dimension.

Índice

1	Introdução	1
1.1	Objectivos da tese	1
1.2	A estrutura da tese	2
1.3	Contextualização histórica	3
2	Conceitos básicos em sistemas dinâmicos	13
2.1	Preliminares	13
2.2	Caracterização dos sistemas caóticos	35
2.3	Grandezas que caracterizam os sistemas dinâmicos	47
2.3.1	Expoentes de Lyapunov	47
2.3.2	Entropia	58
2.3.3	Entropia nos Sistemas Dinâmicos	58
3	Séries temporais	69
3.1	Introdução	69
3.1.1	Reconstrução do atractor	70
3.1.2	Cálculo da dimensão	71
3.2	Expoentes de Lyapunov	106
4	Aplicações à Psicologia	123
5	Conclusão	139

Capítulo 1

Introdução

1.1 Objectivos da tese

Pretende-se com este trabalho, efectuar um estudo aprofundado das técnicas da dinâmica não linear no contexto das séries temporais, nomeadamente, os conceitos de dimensão fractal, expoentes de Lyapunov, entropia aproximada e técnicas de identificação de não linearidade de sistemas. Finalmente apresentaremos um resumo das aplicações destas técnicas no contexto da Psicologia Experimental.

Uma parte significativa da matéria prima para os estudos em Psicologia, refere-se a mudanças observadas ao longo do tempo. O desenvolvimento de um ser humano, varia desde uma criança indefesa a um adulto com responsabilidades.

Um terapeuta observa como um cliente inicialmente ansioso se torna mais confiante e produtivo ao longo de um processo terapêutico. A um nível mais microscópico, a resposta quantitativa de um eléctrodo inserido numa célula da retina, a um estímulo visual, é uma sequência dinâmica de impulsos eléctricos. Na encefalografia, o aparelho, ligado a uma quantidade de eléctrodos colados ao escalpe de um paciente, regista a actividade colectiva de um largo número de processos neurológicos, como um processo dinâmico

multidimensional.

Todas estas sucessões de observações ordenadas no tempo representam comportamentos com dinâmicas complexas.

Esta informação pode então ser usada para compreender, e possivelmente prever, o comportamento.

Na Psicologia são usadas frequentemente medidas de localização da Estatística Descritiva, tais como a média e a variância, para comparar tempos de reacção enquanto variáveis experimentais, tais como a intensidade de estímulos. Acontece que estes métodos fornecem informação insuficiente para determinar os mecanismos dinâmicos envolvidos.

As técnicas de dinâmica não linear fornecem-nos uma grandeza de índices quantitativos que servem para descrever as propriedades dinâmicas do sistema. Estes índices têm sido intensivamente usados nos últimos anos em aplicações práticas, tais como análise de electroencefalogramas, estudo de tempos de reacção e monitorização de comportamentos.

As grandezas mais usadas, incluem a dimensão fractal, que indica a complexidade de informação contida na série temporal, os expoentes de Lyapunov, que indicam a rapidez com que pontos arbitrariamente próximos no espaço de fases, divergem ao longo do tempo, ou a entropia aproximada, que mede o grau de imprevisibilidade de uma série temporal.

1.2 A estrutura da tese

O trabalho está dividido em 5 capítulos. No primeiro capítulo é apresentada a motivação para o uso de séries de dados experimentais na análise não linear e é feita uma breve contextualização histórica.

O capítulo 2 mostra as possíveis origens dos efeitos não lineares e as técnicas utilizadas

recentemente para a análise de fenômenos não lineares. Trata dos conceitos básicos em dinâmica não linear. Estuda as características dos sistemas caóticos e algumas grandezas que caracterizam os sistemas dinâmicos, Expoentes de Lyapunov e Entropia.

No capítulo 3 é apresentado um método que consiste em reconstruir o atrator estranho a partir da série temporal e a partir do atrator reconstruído determinar os expoentes de Lyapunov. Irá também ser estudada a dimensão de correlação D_C , introduzida por Grassberger e Procaccia como uma medida característica de atratores estranhos que serve para distinguir entre o caos determinístico e ruído branco.

O capítulo 4 resume os resultados conhecidos de aplicações à Psicologia.

No capítulo 5 são apresentadas as conclusões deste trabalho.

1.3 Contextualização histórica

A ciência, com as características que lhe conhecemos hoje, nasce com Galileu Galilei e Isaac Newton, como um resultado de uma nova atitude na tentativa para compreender o mundo, baseada na observação cuidadosa dos fenômenos, na elaboração de teorias quantitativas e no teste rigoroso dessas teorias. Para Galileu o mundo das observações simples dos aristotélicos não podia ser o ponto de partida para a compreensão do mundo. Com Galileu a matemática, a par da experiência, assume um papel importante na descrição e compreensão da natureza. Mas, para isso ser possível, houve que inventar uma matemática nova, o cálculo diferencial e integral (Newton e Leibniz), baseada no conceito de derivada. Esta nova ferramenta matemática adequa-se à descrição quantitativa de fenômenos que evoluem no tempo; com ela as leis físicas passam a ser reformuladas como equações diferenciais. Os métodos matemáticos da mecânica são progressivamente aperfeiçoados nos

séculos *XVIII* e *XIX* (Lagrange, Laplace e Hamilton, entre outros).

No final do século *XIX*, o programa científico assente na formulação de teorias físicas com base na experiência e na resolução das equações diferenciais parecia conter uma capacidade ilimitada de previsão e controlo dos fenómenos físicos. Considerava-se que, conhecendo a lei que governa um sistema e as condições iniciais, mesmo que aproximadas, era possível calcular o comportamento do sistema em qualquer instante posterior. O determinismo triunfava. Apesar de se saber que a capacidade de previsão dependia da solução de equações diferenciais e de apenas se saber determinar as soluções das equações mais simples, pensava-se que as soluções conhecidas eram típicas e que os outros casos podiam ser resolvidos aproximadamente, com a precisão desejada, à custa de cálculo. A mecânica torna-se a ciência por excelência e influencia todos os ramos do saber. O espaço e o tempo são absolutos. Num mundo governado pela ordem, não havia lugar para a incerteza permanente e para o caos.

No fim do século *XIX*, no seio da mecânica, considerava-se que restavam “pequenos” problemas. Porém, a sua resolução veio abalar o quadro conceptual Newtoniano, originando novas áreas: a relatividade, a mecânica quântica, a física estatística surgia para interpretar o comportamento de sistemas constituídos por um grande número de partículas - as propriedades macroscópicas não eram senão propriedades médias, calculadas com recurso às probabilidades. Porém, os métodos matemáticos utilizados nos novos domínios, continuam a ser do mesmo tipo: as soluções obtêm-se pela integração de equações diferenciais.

Representando a resolução das equações diferenciais de um dos problemas teóricos mais importantes tanto em matemática como em física, Henri Poincaré começa por adoptar

uma perspectiva qualitativa procurando conhecer os resultados globais para as soluções das equações; o estudo quantitativo será realizado num segundo momento. A topologia permite-lhe franquear este universo totalmente inexplorado, o do qualitativo, conseguindo esclarecer questões até ao momento inextricáveis. Assim, para o caso de um sistema com três graus de liberdade, em vez de estudar uma trajectória, estuda as intersecções sucessivas com um plano (no caso geral, uma superfície) convenientemente escolhida. Se uma trajectória é periódica, a sua intersecção com o plano reduz-se a um ponto (ou um conjunto finito de pontos), caso contrário, os pontos das sucessivas intersecções poderão preencher uma região do plano de maneira aparentemente aleatória. A aplicação do plano no plano, que associa a cada ponto de intersecção da trajectória com o plano, o seguinte, permite estudar a evolução de um sistema a longo prazo, iterando um grande número de vezes esta transformação. O tempo deixa de se considerar contínuo, passa a discreto.

Um dos feitos mais importantes de Poincaré consistiu em estudar os grandes problemas clássicos através de uma abordagem fundada sobre a intuição geométrica do problema. Ao fazê-lo descobre que as soluções não são regulares e periódicas como se supunha. O comportamento geral de um corpo sob a acção de outros dois mais massivos é irregular e imprevisível. Em 1908, em *Science et Méthode*, escreve: "...pode acontecer que pequenas diferenças nas condições iniciais provoquem alterações significativas na evolução do sistema; um pequeno erro nas primeiras produzirá um erro enorme nas últimas. A previsão torna-se impossível... ." (Cit. por Chabert & Dalmedico, 1991, p.567). Sistemas governados por equações diferenciais deterministas podem apresentar um comportamento imprevisível. Poincaré acabava de identificar os "ingredientes" do caos determinista e intuir a sua generalidade.

Nos anos 60 do século *XX* o trabalho dos matemáticos Kolmogorov, Arnold e Moser, permitiu precisar os trabalhos de Poincaré. Inúmeros trabalhos se lhes seguiram assistindo-se, no presente, a um desenvolvimento enorme da investigação matemática nesta área.

Desde o início, Poincaré concebeu o estudo qualitativo das equações diferenciais relacionadas com a questão da estabilidade. No caso de uma evolução periódica instável, Poincaré procurou estabelecer a existência de soluções duplamente assintóticas (no passado e no futuro). Neste caso as soluções cruzam uma infinidade de vezes o plano de Poincaré, dando origem a uma figura extremamente complexa, a qual é estável para períodos de tempo suficientemente grandes e traduz um comportamento complexo, ou “caótico” da dinâmica do sistema. Ruelle e Takens denominaram-nas atratores estranhos e consideraram dever esperar-se que numerosos sistemas dinâmicos apresentem tais atratores, associados ao comportamento caótico do sistema dinâmico.

A existência de um atrator estranho em fenómenos naturais foi posta em evidência por Lorenz, antes de o termo ser inventado, quando aplicava um modelo simplificado ao estudo da circulação atmosférica, representado por um sistema de três equações diferenciais. Deu-se conta que sistemas deterministas podem apresentar sensibilidade à escolha das condições iniciais, desta forma torna-se difícil prever qual o comportamento do sistema num tempo relativamente curto. Apesar disso, o “atrator estranho” e as suas propriedades são independentes das condições iniciais e caracterizam o sistema.

Efectivamente, o século *XX* inicia-se com a intuição e propostas de Poincaré a partir da demonstração de que os modelos lineares usados não poderão nunca dar uma ideia exacta do comportamento a longo prazo das órbitas planetárias, e que o modelo não linear pode esconder trajectórias de uma complexidade insuspeita até então. No terminar

do século *XX* a teoria do caos fornece uma confirmação incontestada da intuição de Poincaré. O conhecimento actual dos modelos não lineares apoia-se, de ora avante, numa multiplicidade de resultados experimentais, simulações numéricas e teoremas matemáticos. Mas, tudo isto só foi possível devido à revolução informática que permitiu realizar cálculos e representações gráficas impossíveis de outra forma.

O mundo tornou-se povoado de fenómenos caóticos, que desde sempre existiram, mas que não eram identificados. Sistemas simples, para determinados valores dos parâmetros exibem comportamentos caóticos ou complexos. É esta matemática que abre novas perspectivas na abordagem em todas as áreas do conhecimento humano.

São poucas as áreas científicas que não se vêem hoje confrontadas com o estudo de fenómenos evoluindo com o tempo. A descrição, explicação, controlo ou previsão de fenómenos como, entre outros, a evolução de um índice bolsista, a variação da inflação ou da taxa de desemprego, a variação da temperatura ou da pluviosidade numa determinada região do globo, a variação da população (segundo a idade, o sexo, ...) num certo país, o movimento de partículas sujeitas a impactos aleatórios num fluido ou a emissão de fótons gerados em desintegração radioactiva têm vindo a interessar especialistas de áreas tão diversas como a Economia, as Ciências Exactas, as Ciências Experimentais e as Ciências Humanas.

Originalmente, as análises estatísticas diziam respeito à descrição de conjuntos numerosos de dados, designados genericamente por informação. Com o nascimento das probabilidades, ao longo dos séculos *XVII*, *XVIII* e *XIX*, foram dados os primeiros passos no domínio da Estatística Matemática cujo objectivo último é o de conhecer o modelo aleatório subjacente a tais observações. Os probabilistas construíram uma impor-

tante teoria matemática, a teoria dos processos estocásticos, a qual permite analisar as propriedades de um conjunto de variáveis indexadas pelo tempo.

A ideia natural de repetição de observações independentes e idênticas, presente na Estatística clássica, não é compatível, em geral, com fenómenos evoluindo com o tempo. De facto, a hipótese de independência das observações perde sentido neste contexto pois estamos perante variáveis cujo comportamento pode depender de um segmento de observações entre dois instantes conhecidos.

De facto, uma das principais questões que se levantam aquando da análise de um fenómeno que evolui ao longo do tempo é o da previsão de valores futuros. Tais estudos só se tornam possíveis se a evolução não for totalmente anárquica.

Assim, a hipótese de independência da estatística clássica é, em particular, substituída por hipóteses de estabilidade no tempo ou de precisão da tendência da evolução (linear, quadrática, crescente ou decrescente, ...).

O estudo da evolução ao longo do tempo, das variáveis que descrevem tais fenómenos pressupõe a obtenção de um certo número de observações dessas mesmas variáveis. Tais observações são obtidas, por exemplo, por meio de inquéritos ou registos. Em domínios como a meteorologia ou a física, certas características são observadas por meio de aparelhos físicos. Nestes casos, o intervalo de tempo no qual o fenómeno se desenrola percorre um intervalo real.

Noutros domínios as observações são recolhidas em tempos geralmente equidistantes, isto é, estão disponíveis por exemplo uma vez por mês, uma vez por trimestre, etc. Identificamos então as datas de observação com os inteiros $1, 2, \dots, T$.

A sucessão de observações de uma variável é chamada *série temporal* ou *série cronoló-*

gica.

Uma série temporal é, assim, um conjunto de dados evoluindo com o tempo onde a ordem de recolha desempenha um papel importante. Este facto é salientado quando se utilizam representações gráficas das séries observadas.

A representação gráfica de uma série temporal é um auxiliar precioso e muito utilizado na análise da sua evolução.

Os primeiros registos conhecidos de representações gráficas de observações temporais estão ligados à Astronomia, em particular à inclinação das órbitas dos planetas em função do tempo, e datam do século *X* ou *XI*.

De seguida vamos introduzir alguns conceitos importantes de dinâmica não linear.

Um sistema determinístico não linear consiste em elementos que têm regras não lineares entre si. O sistema é dinâmico se evolui com o tempo. Se o estado de um sistema dinâmico em certo momento pode ser descrito por M variáveis, este estado pode ser representado por um ponto no espaço de fases de dimensão M .

A evolução no tempo consiste numa série de pontos que formam uma trajectória no espaço dos estados. Quando o tempo tende para infinito, a trajectória ocupará um subespaço do espaço de estados, denominado atractor. O atractor é uma representação geométrica da dinâmica do sistema ao longo do tempo.

Os atractores variam na sua complexidade desde pontos fixos até atractores caóticos. Os atractores são estruturalmente estáveis. Atractores caóticos são estáveis no sentido global, mas instáveis no sentido local. O mecanismo da bifurcação pode transformar atractores que vão desde pontos fixos até um regime caótico.

O funcionamento caótico tem sido encontrado nos sistemas biológicos humanos, no

sistema respiratório, na actividade cardíaca, na actividade cerebral e nas variações hormonais. Por sua vez os Sistemas Dinâmicos dão uma maior flexibilidade de resposta aos diferentes sistemas, fornecendo uma adaptação com as mudanças de ambiente, graças à sua sensibilidade às condições iniciais. Nos trabalhos de Skarda e Freeman, Freeman, Lutzenberger, Nan e Jinghua, Serra e Zanarini e Pritchard indicaram actividade cerebral caótica. Em psicopatologia, os trabalhos de Babloyantz e Destexhe e Reshaw encontraram também actividade caótica de baixa dimensão e alta sintonia de ritmicidade em patologias como demência, esquizofrenia e epilepsia em contraste com sujeitos sãos que apresentam uma actividade de EEG (electroencefalograma) com maior dimensionalidade e maior variabilidade do sinal.

Actualmente, os sistemas dinâmicos têm-se mostrado uma das mais importantes áreas da Matemática que se aplicam em vários ramos da ciência, em particular na Psicologia.

O objectivo de aplicarmos as técnicas da dinâmica não linear é realizar um estudo tanto teórico quanto computacional, da complexidade que um sistema desse tipo pode exibir, tudo isto medido através de conceitos como a entropia, expoentes de Lyapunov e dimensão de correlação.

A teoria dos Sistemas Dinâmicos, como qualquer novo paradigma na ciência, implica novas formas de elaborar questões e analisar fenómenos, novos temas para investigar, e novas regras para determinar o que sabemos e o que pensamos acerca de um fenómeno.

Segundo [20], a teoria do caos deu uma contribuição fundamental a este novo conhecimento. O novo paradigma tem diferentes implicações. Primeiro, muitos fenómenos podem ser vistos com alguma vantagem heurística como fenómenos autoregressivos (variáveis que dependem dos seus próprios valores de um ponto num tempo anterior) que são controlados

em vez de causais. Segundo, as mudanças no comportamento do sistema são governadas pela quantidade de entropia do sistema, pela configuração do atrator e pelas forças repulsoras. Terceiro, muitos fenómenos que apareceram bem fundamentados através de um paradigma mais antigo podem agora ser reexaminados. O nosso objectivo é descrever a forma como se pode aplicar Sistemas Dinâmicos não lineares no estudo de um interessante fenómeno social.

Um índice analítico importante de experiências relacionadas com o caos é o cálculo do expoente de Lyapunov de um atrator.

O cálculo da dimensão fornece-nos um teste de hipóteses, se um atrator é um ponto fixo ou um ciclo-limite, e por conseguinte, não caótico, neste caso o expoente é negativo, ou se o atrator é aperiódico ou caótico, sendo neste caso, o expoente positivo.

A ideia de estimar a dimensão por métodos estatísticos é um assunto de controvérsia, que têm a ver com o facto, de alguns autores acharem que caos determinístico e estatística nunca poderem ter qualquer terreno comum.

Historicamente, houve uma tendência de evitar uma relação entre sistemas dinâmicos não lineares e estatística. O livro de Stewart começa com uma história do desenvolvimento paradigmático na física, incluindo a era estatística, que mais tarde foi substituída pelos sistemas dinâmicos não lineares, na qual usou equações diferenciais como ferramenta analítica primária nas ciências naturais. Casti deu mais um passo alegando que a estatística estava a tornar-se um método obsoleto para testar e desenvolver a teoria, um ponto no qual Guastello em [20] não concorda. Peters deu uma visão mais equilibrada referindo que os métodos não estatísticos para estimar os expoentes de Lyapunov não são especialmente confiáveis para alguns tipos de dados. E Priesmeyer referiu que a estatística tradicional

não é adequada para descrever fenómenos dado que a sua dependência dos modelos lineares descarta desvios da linearidade como erros. Portanto, não devemos perder de vista a palavra-chave *tradicional*.

Existem aplicações dos sistemas dinâmicos para um amplo espectro de teorias em psicologia cognitiva. Podemos começar com a psicofísica, que relaciona estímulos físicos e a percepção psicológica. O paradigma da experiência convencional tem sido baseado em tempos de reacção, ou no atraso temporal entre o estímulo presente ou um padrão e cada padrão de reconhecimento ou tomada de uma decisão relacionada com o conteúdo desses padrões. O tempo de reacção é objecto de muitas influências, e estudos de tempos de reacção são pertinentes particularmente na engenharia, nos factores humanos e nos fluxos de informação entre as pessoas e as suas máquinas.

Capítulo 2

Conceitos básicos em sistemas dinâmicos

2.1 Preliminares

A dinâmica diz respeito ao estudo do comportamento dos sistemas que evoluem no tempo. Um sistema dinâmico determinista designa o processo de evolução temporal e é constituído por um conjunto de funções (regras e equações) que especificam como é que as variáveis de estado mudam com o tempo. O estado de um sistema dinâmico é especificado pelo valor das suas variáveis em cada instante.

É importante distinguir variáveis de parâmetros, as variáveis mudam com o tempo, os parâmetros que caracterizam um sistema não. As variáveis discretas tomam valores num conjunto discreto, as contínuas num conjunto contínuo. Quando falamos de variáveis discretas falamos de sistemas dinâmicos discretos, que são de extrema importância para o tema desta tese.

O objectivo básico da teoria dos sistemas dinâmicos é compreender o comportamento eventual ou assintótico de um processo iterativo. Se este processo é uma equação diferencial cuja variável independente é o tempo, então a teoria dos sistemas dinâmicos fornece-

nos uma previsão do comportamento das soluções da equação mesmo num futuro mais distante ($t \rightarrow \infty$) ou num passado distante ($t \rightarrow -\infty$). Se o processo é um processo discreto, tal como a iteração de uma função, a teoria dos sistemas dinâmicos pretende compreender o comportamento eventual dos pontos $x, f(x), f(f(x)), f(f(f(x))) \dots$ quando o número de iteradas se torna cada vez maior. Para melhor compreender estes conceitos vamos introduzir algumas definições.

Chamamos órbita de x segundo f ao conjunto de todas as iteradas de f em x , ou seja, ao conjunto $\{x, f(x), f(f(x)), \dots\}$.

Para compreender a evolução do sistema, começando numa dada condição inicial x , podemos colocar num gráfico a órbita de x e observar o que acontece após um grande número de iterações. Podemos fazê-lo de duas formas diferentes: através do gráfico da série temporal ou através da iteração gráfica. O procedimento para o gráfico da série temporal consiste em colocar no eixo horizontal o número de iteradas e no eixo vertical, o valor das iteradas. Assim, a condição inicial é sempre o ponto de ordem 0 da órbita. A iterada seguinte é o primeiro ponto da órbita e assim sucessivamente. Começamos por calcular o valor da função numa condição inicial x_0 e vamos obtendo:

$$\begin{aligned} &\text{condição inicial } x_0 \\ &1^{\text{a}} \text{ iterada } x_1 = f(x_0) \\ &2^{\text{a}} \text{ iterada } x_2 = f(x_1) = f(f(x_0)) \\ &\dots \\ &n\text{-ésima iterada } x_n = \underbrace{f \circ f \dots \circ f}_{n \text{ vezes}}(x_0) = f^n(x_0). \end{aligned}$$

De seguida, formamos os pares ordenados $(0, x_0), (1, x_1), (2, x_2), \dots$ e representamo-los num referencial. Repare que o valor no eixo dos xx representa o tempo (discreto) e o valor no eixo dos yy representa o valor das iteradas. Resulta o seguinte gráfico:

Para a iteração gráfica o procedimento consiste em colocar primeiramente, no mesmo

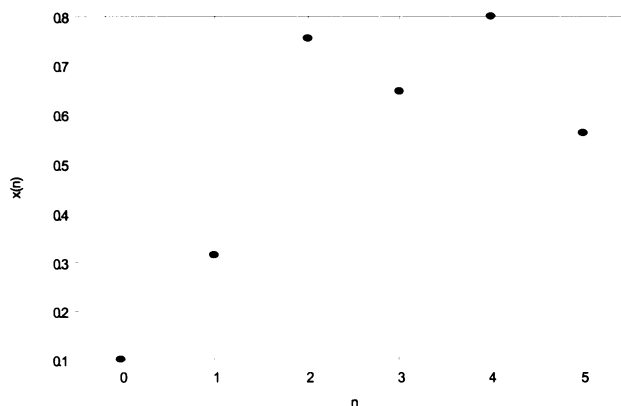


Figura 2.1: Gráfico da série temporal da função $f(x) = 3.5x(1-x)$ e $x_0 = 0.1$.

sistema de eixos coordenados, o gráfico da função que estamos a iterar e a diagonal $y = x$. De seguida representamos o itinerário do ponto x_0 seguindo graficamente o movimento desse ponto. Começamos em (x_0, x_0) na diagonal e deste ponto, seguimos verticalmente até alcançar o gráfico de f . O valor de x do ponto de intersecção deve ser x_0 e o valor de y deve ser $f(x_0)$. Do ponto $(x_0, f(x_0))$, seguindo horizontalmente para a diagonal, chegamos ao ponto $(f(x_0), f(x_0))$. Repetindo o processo, seguimos verticalmente para o gráfico e chegamos ao ponto $(f(x_0), f^2(x_0))$ e de seguida horizontalmente até $(f^2(x_0), f^2(x_0))$ na diagonal. Este procedimento deverá prosseguir, sempre do gráfico da função para a diagonal, horizontalmente e da diagonal para o gráfico da função verticalmente. Como se observa na figura 2.2.

Como veremos, de seguida, os pontos fixos são importantes na classificação das dinâmicas.

Definição 2.1.1 *Se f é uma função e $f(c) = c$, então c é um ponto fixo de f .*

Uma função tem um ponto fixo c se e só se o seu gráfico intersecta a diagonal $y = x$

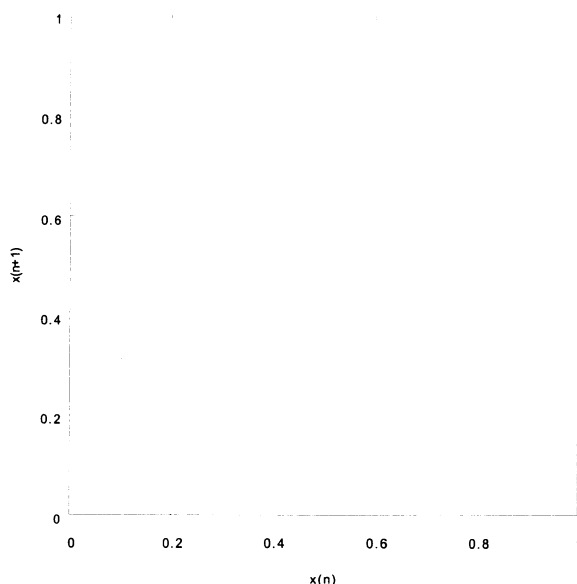


Figura 2.2: Iteração gráfica para a função $f(x) = 3.5x(1-x)$ e $x_0 = 0.1$.

no ponto (c, c) .

Teorema 2.1.1 *Seja $I = [a, b]$ um intervalo fechado e $f : I \rightarrow I$ uma função contínua. Então, f tem um ponto fixo em I .*

Demonstração: Seja $I = [a, b]$ e $f : I \rightarrow I$ contínua. Se $f(a) = a$ ou $f(b) = b$, logo um deles a ou b é ponto fixo. Suponhamos que $f(a) \neq a$ e $f(b) \neq b$. Seja $g(x) = f(x) - x$, g é contínua pois é a diferença de duas funções contínuas. Como $f(a) \neq a$ e $f(a)$ está no intervalo $[a, b]$, $f(a) > a$. Do mesmo modo, $f(b) < b$. Logo, $g(a) = f(a) - a > 0$ e $g(b) = f(b) - b < 0$. Como g é contínua, pelo Teorema do Valor Intermédio implica que existe um c em $[a, b]$ tal que $g(c) = 0$. Mas $g(c) = f(c) - c = 0$ então $f(c) = c$.

Teorema 2.1.2 *Seja I um intervalo fechado e $f : I \rightarrow \mathbb{R}$ uma função contínua. Se $f(I) \supset I$, então f tem um ponto fixo em I .*

Demonstração: Seja $I = [a, b]$. Como $f(I) \supset I$ existe um c e um d em I tal que $f(c) = a$ e $f(d) = b$. Se $c = a$ ou $d = b$, fica demonstrado. Se não, temos $a < c < b$ e $a < d < b$. Se definirmos $g(x) = f(x) - x$, logo, $g(c) = f(c) - c < 0$ pois $f(c) = a$ e $a < c$. Caso contrário, $g(d) = f(d) - d > 0$. Como $g(c) < 0$ e $g(d) > 0$ e g é uma função contínua, pelo Teorema do Valor Intermédio, existe um e entre c e d (e em I) que satisfaz $g(e) = 0$ e $f(e) = e$, o que completa a demonstração.

Há vários exemplos de funções onde podemos observar estes tipos de comportamento. O modelo de crescimento populacional, estudado por Verhulst (1845) e reinterpretado por May (1976) é um exemplo usualmente apresentado de comportamento linear.

Consideremos uma espécie de insectos em que os indivíduos nascem e morrem na mesma estação. Poderemos saber qual o número de insectos no ano seguinte. Numa situação simplista esse número será dado por:

$$x_{n+1} = p x_n \quad (2.1)$$

onde p é um parâmetro que depende das condições ambientais (disponibilidade de alimento, de água, etc.), x_n a população num dado ano e x_{n+1} no ano seguinte. Se $p > 1$, o número de insectos crescerá; se $p < 1$, o número de insectos decrescerá, tendendo para a extinção; se $p = 1$, permanecerá constante.

Exemplo 2.1.1 Consideremos, em particular, $p = 0.5$. Consideramos então a função

$$f(x) = 0.5 x$$

para duas condições iniciais diferentes, $x_0 = 2$ e $x_0 = -2$. Construimos os gráficos das séries temporais e da iteração gráfica da função para as duas condições dadas.

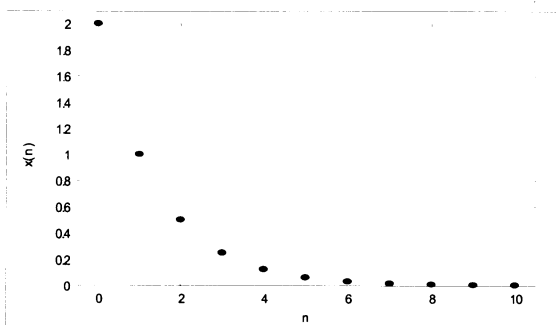
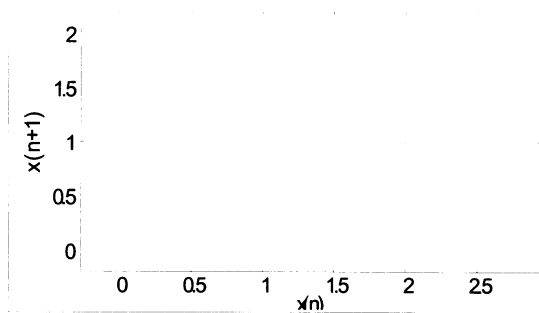
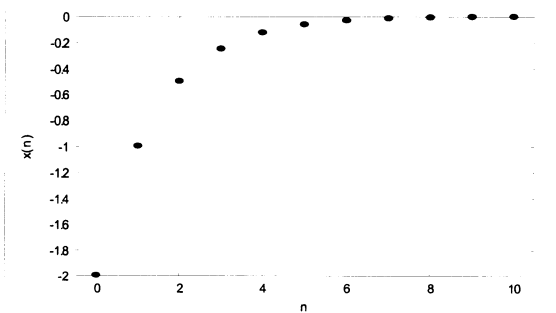
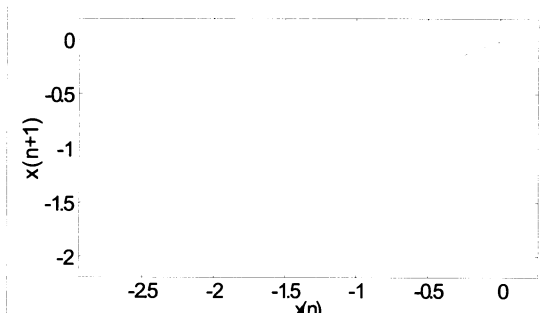
a) $x_0 = 2$.b) $x_0 = 2$.c) $x_0 = -2$.d) $x_0 = -2$.

Figura 2.3: Os gráficos a) e c) são as séries temporais da função $f(x) = 0.5x$ para as condições iniciais $x_0 = 2$ e $x_0 = -2$, respectivamente. Os gráficos b) e d) são as iterações gráficas da mesma função e da mesma condição inicial da série temporal do lado esquerdo.

Como a função $f(x) = 0.5x$ intersecta a diagonal $f(x) = x$ no ponto $(0, 0)$, temos que 0 é único ponto fixo da função.

Estes gráficos indicam-nos o comportamento deste sistema para duas condições iniciais diferentes, o 2 e o -2 através da série temporal e através de iteração da função. Verificamos que $f^n(2)$ e $f^n(-2)$ aproximam-se de 0 à medida que aumentamos o número de iteradas.

De seguida, vamos observar o comportamento de uma função não linear.

Exemplo 2.1.2 Consideremos a função $f(x) = x^2$.

Temos 0 e 1 como pontos fixos, pois $0^2 = 0$ e $1^2 = 1$.

Se considerarmos outra condição inicial diferente de 0 e de 1, por exemplo, $x_0 = 2$, obtemos a órbita seguinte: $\{2, 4, 16, 128, 16384, \dots\}$. Vemos que a órbita tende para infinito. Obtínhamos o mesmo comportamento se a condição inicial fosse 1.2. É possível verificar que para qualquer condição no intervalo $]-\infty, -1[\cup]1, +\infty[$, a órbita tende para infinito.

Consideremos agora a condição inicial igual a 0.5 obtemos a órbita $\{\frac{1}{2}, \frac{1}{4}, \frac{1}{256}, \frac{1}{16384}, \dots\}$. Obtemos números cada vez mais pequenos. Podemos concluir que, para qualquer condição inicial no intervalo $]-1, +1[$, a órbita tende para zero.

Podemos observar assim três tipos diferentes de comportamentos nas órbitas, dependendo da condição inicial tomada.

Definição 2.1.2 Sejam f uma função e k um inteiro positivo. O ponto x é ponto periódico de f com período k se $f^k(x) = x$. Por outras palavras, um ponto é um ponto periódico de f com período k se é ponto fixo de f^k . O ponto periódico x tem um período k_0 primitivo se $f^{k_0}(x) = x$ e $f^n(x) \neq x$ quando $0 < n < k_0$. Ou seja, um ponto periódico tem período primitivo k_0 se ele retorna ao ponto inicial, pela primeira vez, depois de iterarmos

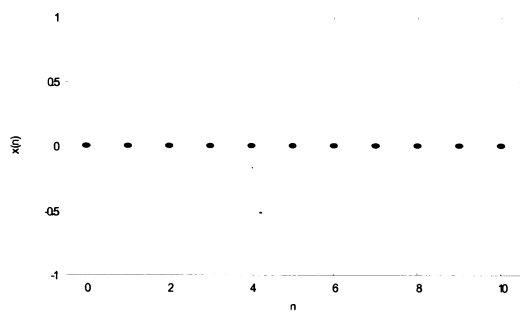
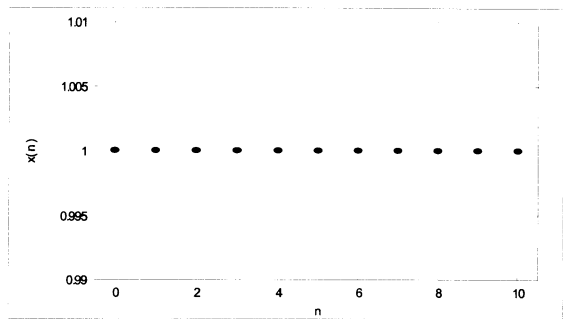
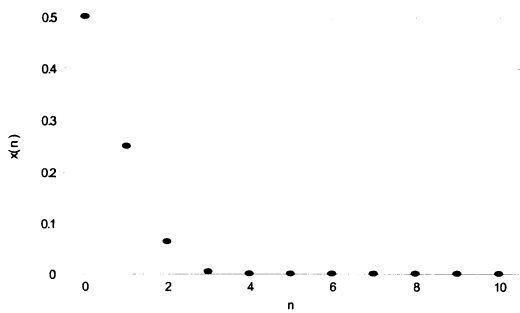
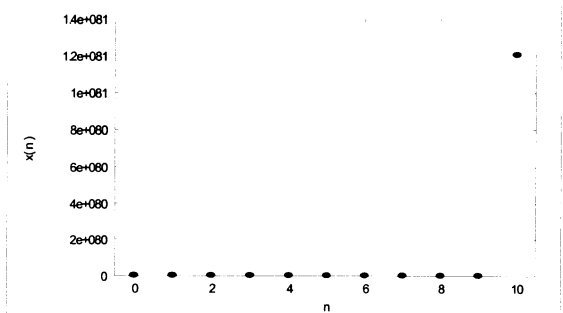
a) $x_0 = 0$.b) $x_0 = 1$.c) $x_0 = 0.5$.d) $x_0 = 1.2$.

Figura 2.4: Nos gráficos a), b), c) e d) observamos os comportamentos das órbitas para a função $f(x) = x^2$, nas condições iniciais $x_0 = 0$, $x_0 = 1$, $x_0 = 0.5$ e $x_0 = 1.2$, respectivamente.

exactamente k_0 iterações de f . Se x é ponto periódico, ao conjunto de todas as iteradas do ponto periódico x damos o nome de órbita periódica ou ciclo periódico.

Exemplo 2.1.3 Consideremos a função $f(x) = -x^3$. Os pontos -1 e 1 formam uma órbita periódica com período primitivo 2. O conjunto de pontos com período 2 é $\{-1, 1, 0\}$, enquanto o conjunto de pontos com período primitivo 2 é $\{-1, 1\}$. O único ponto fixo é o ponto 0. A órbita de 0 é $\{0\}$ e a órbita de 2 é $\{2, -2^3, 2^9, -2^{27}, \dots\}$.

Uma função pode ter muitos pontos fixos e pontos periódicos. A função $f(x) = x$ fixa qualquer ponto, enquanto qualquer ponto excepto 0 é ponto periódico com período primitivo 2 para $g(x) = -x$.

Definição 2.1.3 O ponto x é eventualmente periódico com período k se existe um inteiro positivo N tal que $f^{n+k}(x) = f^n(x)$ quando $n \geq N$.

Exemplo 2.1.4 Seja $g(x) = |x - 1|$. Então 0 e 1 formam uma órbita periódica e $f(2) = 1$, logo 2 é eventualmente periódico. De facto, qualquer inteiro é eventualmente periódico.

Definição 2.1.4 Seja p um ponto periódico de período n . Um ponto x é futuramente assintótico para p se $\lim_{i \rightarrow \infty} f^{in}(x) = p$.

No início dos anos 70, Robert May e outros matemáticos começaram a investigar talvez um dos sistemas mais famosos usado para modelar espécies simples, conhecido como a função logística dada por:

$$f_a : [0, 1] \longrightarrow [0, 1] \tag{2.2}$$

$$x_{n+1} = f_a(x_n) = ax_n(1 - x_n)$$

O parâmetro a pertence ao intervalo $[0, 4]$ representando a taxa de crescimento da população, x_n a população no ano n , e x_{n+1} a população no ano seguinte. Naturalmente, taxas de crescimento demasiado baixas conduzem à extinção e taxas de crescimento mais elevadas dão lugar a uma população mais numerosa. Neste caso a equação é não linear e o comportamento pode tornar-se imprevisível. Observemos, nos exemplos seguintes, alguns tipos de comportamento possíveis, variando o parâmetro a e mantendo a condição inicial $x_0 = 0.04$ fixa. Nas figuras 2.5 à 2.8, à esquerda está representada a série temporal dos valores de n que se obtêm a partir da iteração gráfica representada à direita para $a = 2.8$, $a = 3.1$, $a = 3.45$ e $a = 3.5699$, respectivamente:

Exemplo 2.1.5 Se $a = 2.8$, x_n estabiliza no valor $18/28 \approx 0.643$. Significa que a população permanecerá em equilíbrio estável.

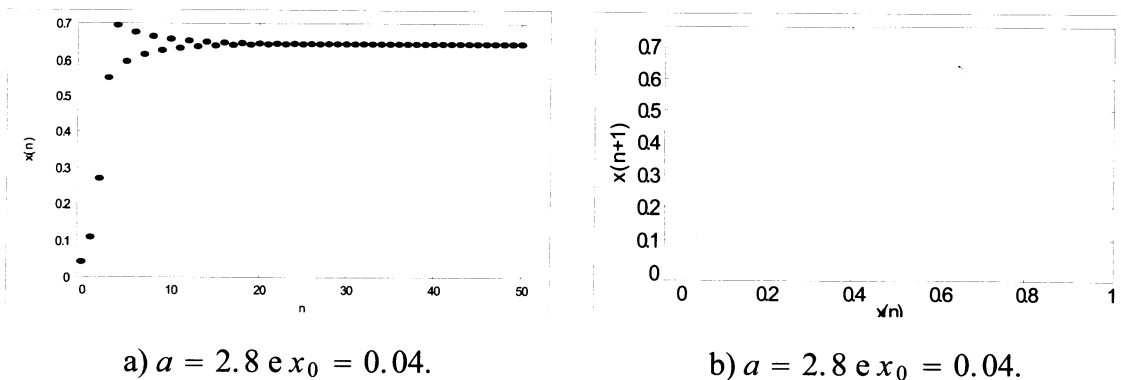


Figura 2.5: O gráfico a) é a série temporal para $a = 2.8$ e para $x_0 = 0.04$ e o gráfico b) é a iteração gráfica para a mesma função e a mesma condição inicial.

Exemplo 2.1.6 Se $a = 3.1$, x_n oscila, assintoticamente, entre dois valores: 0.558 e 0.764. Significa que a população num ano cresce, no outro decresce e assim sucessivamente.

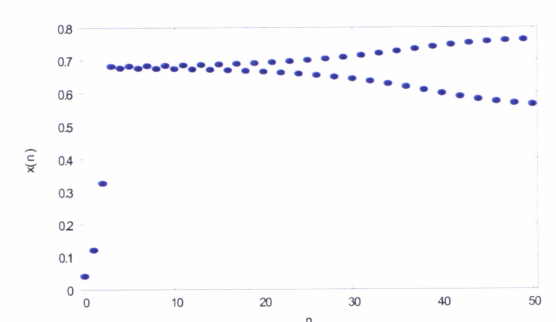
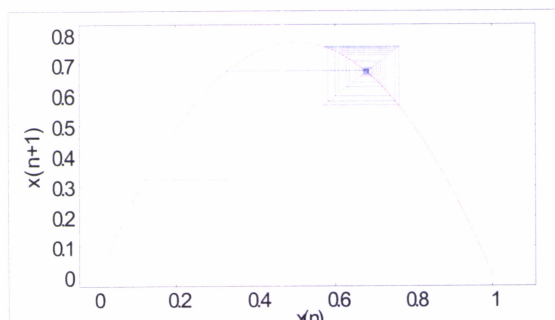
a) $a = 3.1$ e $x_0 = 0.04$.b) $a = 3.1$ e $x_0 = 0.04$.

Figura 2.6: O gráfico a) é a série temporal para $a = 3.1$ e para $x_0 = 0.04$ e o gráfico b) é a iteração gráfica para a mesma função e a mesma condição inicial.

Exemplo 2.1.7 Se $a = 3.45$, x oscila, assintoticamente entre quatro valores: 0.852, 0.433, 0.847, 0.447. Os valores ocorrem sempre pela ordem indicada, o comportamento é assintoticamente periódico, de período 4.

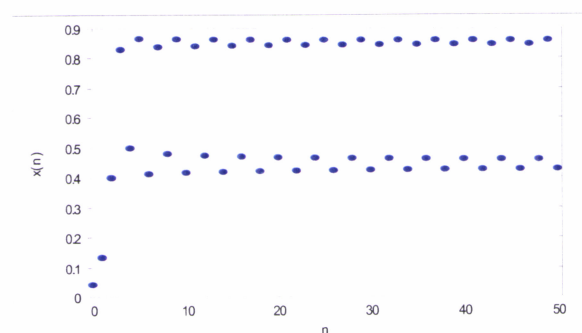
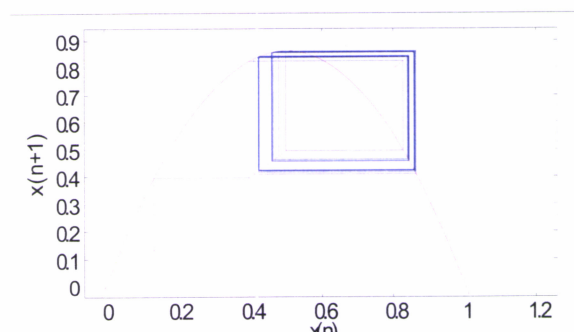
a) $a = 3.45$ e $x_0 = 0.04$.b) $a = 3.45$ e $x_0 = 0.04$.

Figura 2.7: O gráfico a) é a série temporal para $a = 3.45$ e para $x_0 = 0.04$ e o gráfico b) é a iteração gráfica para a mesma função e a mesma condição inicial.

Exemplo 2.1.8 Aumentando o valor de a para $a = 3.5699$ surge um novo comportamento assintoticamente periódico. O comportamento periódico seguinte é o de período 8.

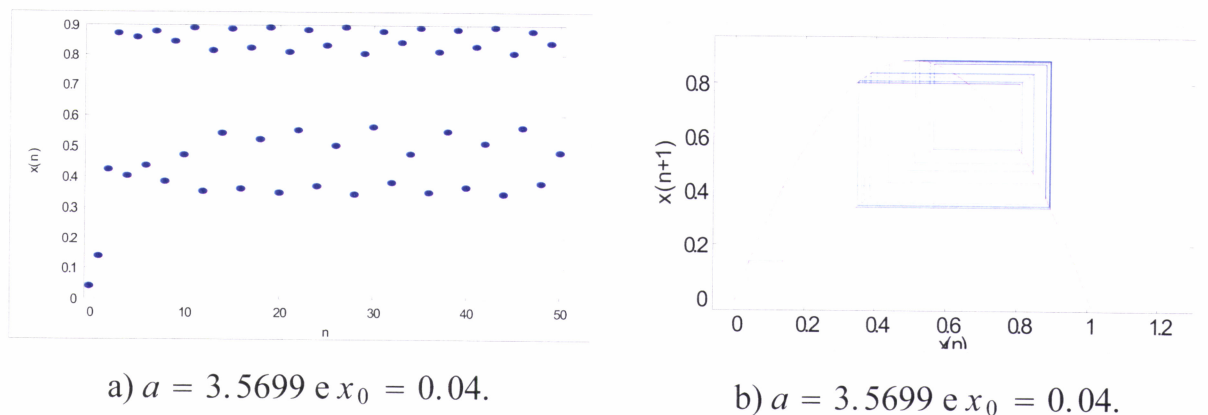


Figura 2.8: O gráfico a) é a série temporal para $a = 3.5699$ e para $x_0 = 0.04$ e o gráfico b) é a iteração gráfica para a mesma função e a mesma condição inicial.

Esta função é aparentemente simples, mas as suas propriedades estão longe da trivialidade. Do ponto de vista matemático, tem a vantagem de ser extremamente simples, mas como modelo populacional tem a vantagem de exigir que x pertença ao intervalo $[0, 1]$ dado que para $x \notin [0, 1]$, a sequência de valores de $f^n(x)$ tende para $-\infty$.

Analisando os gráficos, vemos que a população permanecerá em equilíbrio até o parâmetro permanecer abaixo de um valor do parâmetro. Ultrapassado este valor, o equilíbrio é rompido e a população oscila entre dois valores distintos, depois entre quatro, entre oito, e assim sucessivamente. As bifurcações continuam à medida que a vai aumentando, até que a evolução se torna caótica, passando de forma imprevisível, por uma multiplicidade de valores distintos.

Exemplo 2.1.9 Consideremos a função logística (2.2) com $a = 4$. Temos um ponto fixo em 0. O ponto 1 não é fixo mas, iterando a função uma vez, $f_4(1) = 0$, caímos no ponto fixo. Igualmente, $f_4(0.5) = 1$ e $f_4(0.5) = 0$ logo após duas iterações 0.5 cai em 0.

A existência de termos não lineares faz com que uma pequena variação no parâmetro

possa conduzir a uma variação brusca tanto qualitativa como quantitativa no comportamento do sistema. Para um dado valor, o comportamento pode ser periódico, para outro ligeiramente diferente do primeiro, o comportamento pode ser aperiódico.

Definição 2.1.5 *Seja p um ponto periódico de período primitivo n . O ponto p é hiperbólico se $|(f^n)'(p)| \neq 1$ e não hiperbólico se $|(f^n)'(p)| = 1$. No caso particular, de $n = 1$ chamamos a p ponto fixo hiperbólico.*

Proposição 2.1.1 *Seja p um ponto fixo hiperbólico de uma função $f \in C^1$ com $|f'(p)| < 1$. Então existe um intervalo aberto U contendo p tal que, se $x \in U$, então $\lim_{n \rightarrow \infty} f^n(x) = p$.*

Demonstração: Desde que f seja C^1 , existe $\varepsilon > 0$ tal que $|f'(x)| < A < 1$ para $x \in [p - \varepsilon, p + \varepsilon]$. Pelo Teorema do Valor Médio $|f(x) - p| = |f(x) - f(p)| \leq A|x - p| < |x - p| \leq \varepsilon$. Então $f(x)$ está contido em $[p - \varepsilon, p + \varepsilon]$ e, de facto, mais próximo de p do que x a p . Com a mesma argumentação temos que $|f^n(x) - p| \leq A^n|x - p|$, logo $f^n(x) \rightarrow p$, quando $n \rightarrow \infty$.

Nota: Este resultado também é válido para pontos periódicos de período n . Neste caso, podemos obter um intervalo aberto que contém p que se aplica em si próprio por f^n . Evidentemente, neste caso a hipótese é que $|(f^n)'(p)| < 1$.

Definição 2.1.6 *Um ponto fixo p com $|f'(p)| < 1$ é chamado de ponto fixo atractivo.*

Um ponto periódico p , hiperbólico de período n , com $|(f^n)'(p)| < 1$, é chamado de ponto periódico atractivo.

Proposição 2.1.2 *Seja p um ponto fixo hiperbólico com $|f'(p)| > 1$. Então existe um*

¹ f diferenciável com derivada contínua.



intervalo aberto U contendo p tal que se $x \in U$, $x \neq p$, então existe $k > 0$ tal que $f^k(x) \notin U$.

A demonstração é semelhante à proposição anterior.

Definição 2.1.7 Um ponto fixo p com $|f'(p)| > 1$ é chamado de ponto fixo repulsivo.

Um ponto periódico p , hiperbólico de período n , com $|(f^n)'(p)| > 1$, é chamado de ponto periódico repulsivo.

Nota: Seja p_1 um ponto periódico de período k e $\{p_1, p_2, \dots, p_k\}$ a sua órbita. Pela regra da cadeia temos que:

Nota 2.1.1

$$\begin{aligned} (f^k)'(p_1) &= (f(f^{k-1}))'(p_1) = f'(f^{k-1}(p_1))(f^{k-1})'(p_1) = \\ &= f'(f^{k-1}(p_1))f'(f^{k-2}(p_1))\dots f'(p_1) = f'(p_k)f'(p_{k-1})\dots f'(p_1) \end{aligned}$$

Realçamos a derivada de f^n no ponto p dado que nos auxilia na determinação se p é atractivo ou repulsivo. A derivada de f por si só não nos fornece informação sobre a natureza do ponto periódico.

No caso da derivada ser igual a 1 podemos observar comportamentos diferentes.

Exemplo 2.1.10 Temos que 0 é ponto fixo de $g(x) = x - x^3$ e que atrai pontos próximos.

No entanto, $g'(0) = 1$. De forma semelhante temos que 0 é ponto fixo de $g(x) = x + x^3$ e que repele os pontos próximos, mas $f'(0) = 1$. Mais precisamente, suponhamos que p é ponto periódico de f com período primitivo k , e $|(f^k)'(p)| = 1$. Se o conjunto estável de p contém um intervalo aberto que contém p , então p é ponto periódico atractivo de f .

Se existe um intervalo aberto J que contém p , e que satisfaz para todo o x em J que não p existe um número natural n tal que $f^{nk}(x)$ não está em J , logo p é ponto periódico repulsivo de f .

Não nos é difícil de imaginar que um sistema governado por equações matemáticas complicadas possa ter um comportamento complicado e imprevisível. Mas o que surpreendeu a maioria dos cientistas foi o facto de até sistemas simples, descritos por equações simples, poderem exibir comportamentos caóticos. A ordem e o caos podem aparecer no mesmo sistema.

O termo bifurcação é muitas vezes usado nos estudos de dinâmicas não lineares, e descreve as mudanças do comportamento de um sistema à medida que variamos um parâmetro.

De seguida, consideremos a função logística (2.2) com $0 < a < 4$.

Um diagrama de bifurcação é a representação gráfica do comportamento qualitativo das órbitas para cada valor do parâmetro a . A maneira de o construir é muito simples. Construimos um gráfico em que o eixo horizontal corresponde aos valores de a e o eixo vertical aos valores de x_n . Para cada valor de a , escolhe-se ao acaso uma condição inicial, e gera-se a órbita correspondente. Os primeiros N pontos da órbita são descartados, para dar tempo a que a órbita evolua para o seu comportamento final, quer este seja uma órbita periódica ou caótica - chama-se a isto eliminar o transiente. Em seguida começamos a marcar no gráfico, para esse valor de a , os valores de x_n assumidos pela órbita ao longo de um número M bastante grande de iterações. As iterações produzem valores de x_n num intervalo entre $[0, 1]$ se a condição inicial está nesse intervalo.

No gráfico 2.9 a condição inicial foi 0.5 e descartamos as primeiras $N = 50$ itera-

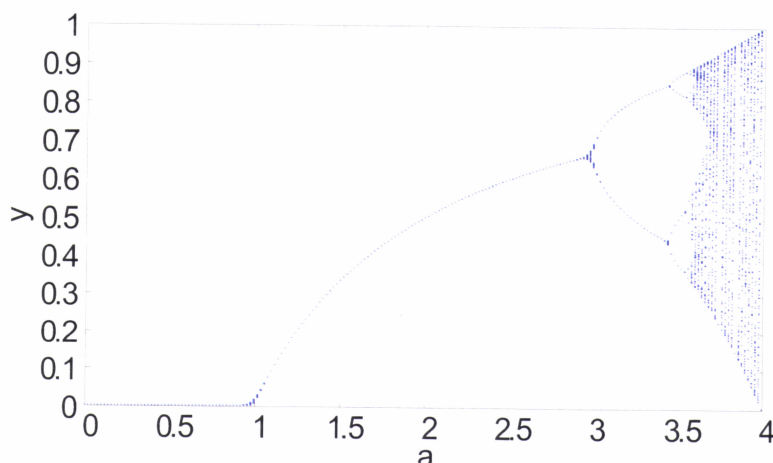


Figura 2.9: Diagrama de Bifurcação da função logística $f_a(x_n) = ax_n(1 - x_n)$, com $0 < a < 4$.

das x_1, x_2, \dots, x_{50} e foram desenhadas no gráfico as iteradas $x_{51}, x_{52}, \dots, x_{100}$. Esta figura mostra-nos uma grande variedade de comportamentos. Desde convergência para os pontos fixos passando por bifurcações de duplicação do período ao caos.

Um aspecto importante para compreendermos os sistemas dinâmicos é a caracterização dos atractores.

Um *Atractor* é um conjunto tal que as trajectórias a partir de quase todas as condições iniciais possíveis acabam por cair nele.

Ao conjunto das condições iniciais que dão origem a trajectórias que se aproximam de um dado atractor chama-se *bacia de atracção* desse atractor.

Um atractor é um conjunto de estados de um sistema para o qual estados vizinhos, a chamada bacia de atracção, convergem. Seja qual for o estado inicial, o sistema evoluirá até chegar ao atractor, e permanecerá nesse estado, na ausência de outros factores.

No caso da logística (2.2), por exemplo, os atractores são diferentes para os vários valores do parâmetro a . Para um dado valor de a a condição inicial escolhida está na bacia

de atracção do ponto fixo, apenas aparecerá um ponto no eixo dos yy . Como podemos ver no diagrama esta situação verifica-se para valores do parâmetro inferiores a três. Para $3 < a < 1 + \sqrt{6}$ a condição inicial está na bacia de atracção de um atractor periódico de período 2. Para a ligeiramente superior a $1 + \sqrt{6} \approx 3.449$, por exemplo, $a = 3.5$, há uma duplicação do período que se tornará de período 4, de seguida de período 8 e vai duplicando sucessivamente. Este comportamento ocorre para subintervalos de a ao longo do diagrama. Para cada intervalo dos valores do parâmetro no qual o único atractor é a órbita periódica chamamos de Janela Periódica. Para $a = 3.73$, as dinâmicas tornam-se muito complicadas, e as iteradas de uma condição inicial movem-se caoticamente no intervalo $[0, 1]$. O caos e a ordem não percorrem caminhos distintos como se poderia pensar, mas estão ligados entre si. Através do diagrama da figura 2.9 podemos observar a transição entre a ordem e o caos. Se fizermos um recorte ampliado do Diagrama de Bifurcação como se observa na figura 2.10 podemos verificar novos pormenores, mas no todo tem o mesmo aspecto que o original.

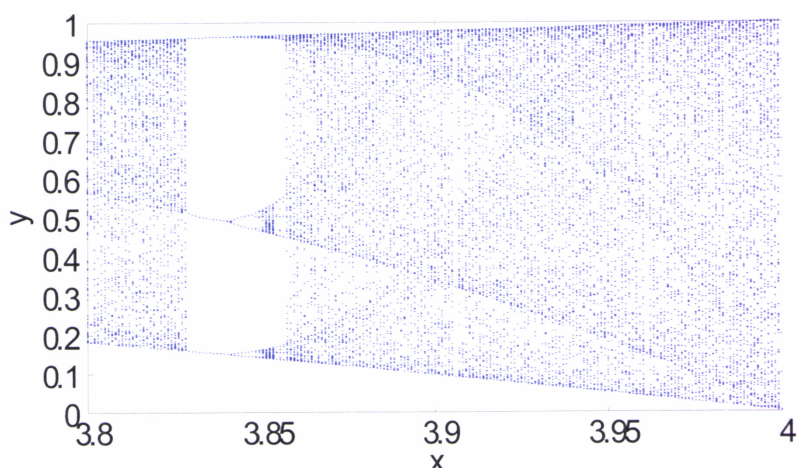


Figura 2.10: Diagrama de Bifurcação da função logística $f_a(x_n) = ax_n(1 - x_n)$, com $3.5 < a < 4$.

Estas formações são invariantes relativamente à escala. Mesmo na parte caótica, escondem-se ramos de estrutura análoga ao todo.

Na década de setenta, o físico Mitchell J. Feigenbaum estudou as propriedades da função logística (2.2) e descobriu uma propriedade interessante acerca da duplicação do período.

Consideremos $a_1, a_2, a_3, \dots, a_k, \dots$ a sequência de valores de a para os quais o sistema bifurca. Os primeiros valores da sequência são: $a_1 = 3.0$, $a_2 = 3.449490\dots$, $a_3 = 3.544090\dots$, $a_4 = 3.564407\dots$, $a_5 = 3.568759\dots$, $a_6 = 3.569692\dots$, $a_7 = 3.569891\dots$. Feigenbaum descobriu que se d_k é definido por $d_k = a_{k+1} - a_k$, então

$$\delta = \lim_{k \rightarrow \infty} \frac{d_k}{d_{k+1}} = 4.669202\dots$$

À razão de convergência δ da sequência (a_k) dá-se o nome de número de Feigenbaum. Este número é a constante de proporcionalidade para a duplicação do período. Além de aparecer em várias funções matemáticas, aparece também em numerosos sistemas físicos, caracterizando a transição entre ordem e caos. Este aspecto é muito importante, na medida em que ilustra a universalidade dos comportamentos caóticos, expressa na constante (universal) de Feigenbaum. Por outro lado, este mecanismo, a cascata de duplicações do período, tornou-se muito importante para distinguir a existência de caos nos fenómenos naturais, isto caso o fenómeno em causa dependa de um parâmetro cuja variação permita encontrar a correspondente sequência de comportamentos. Segundo [37], uma interpretação possível e útil da constante δ é que poderá ser usada para fazer uma previsão. Por exemplo, fazendo uma medição de duas sucessivas bifurcações devemos ser capazes de prever as bifurcações seguintes e até onde vai o seu limite. Esta descoberta foi de uma grande importância e deu-nos um testemunho da beleza e aventura do pensamento matemático.

Até ao momento, abordámos diversos comportamentos possíveis em funções de dimensão um.

Na realidade, não é usual vermos um número pequeno de elementos a interagirem. Na maioria dos sistemas, há uma rede complexa de interacções. Nestas circunstâncias, têm se desenvolvido e formulado modelos teóricos com equações diferenciais lineares e não lineares com duas ou mais variáveis.

De forma a compreendermos os diversos tipos de dinâmicas que podemos encontrar em equações diferenciais com mais de duas variáveis, vamos considerar as equações diferenciais com três variáveis, inicialmente, estudadas pelo meteorologista E.N. Lorenz.

O sistema de Lorenz foi introduzido, em 1963, como um modelo simples do movimento convectivo nas camadas superiores da atmosfera.

O comportamento do modelo de Lorenz é representado traçando as suas variáveis no espaço de estados, ou seja, para cada cálculo sucessivo de x , y , e z , traçamos o ponto correspondente num espaço de eixos xyz .

A determinação das trajectórias a partir das condições iniciais no espaço de fases permite conhecer a evolução do sistema.

Até à década de 60, os atractores eram pensados como sendo subconjuntos geométricos muito simples do espaço de fases: pontos, linhas, superfícies e volumes. Como exemplo de dois atractores clássicos simples temos o ponto fixo, caso do pêndulo amortecido, e o ciclo-limite, caso do chamado oscilador de Van der Pol.

O sistema de Lorenz é habitualmente descrito como um sistema de três equações dife-

renciais não lineares, dependendo de três parâmetros s , r e b .

$$dx/dt = s(y - x) \quad (2.3)$$

$$dy/dt = rx - y - xz$$

$$dz/dt = xy - bz$$

A figura 2.11 mostra o atrator de Lorenz para os parâmetros $s = 10$, $b = \frac{8}{3}$ e $r = 24.4$.

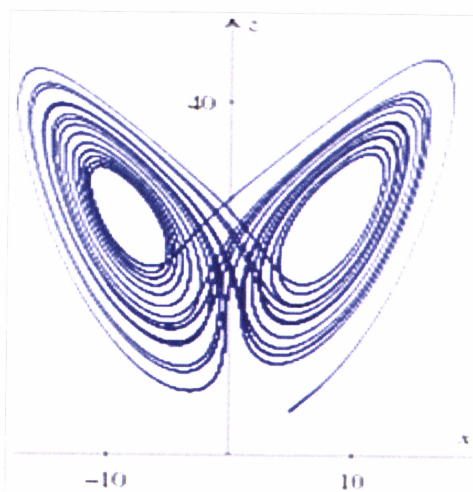


Figura 2.11: Atrator de Lorenz para os parâmetros $s = 10$, $b = \frac{8}{3}$ e $r = 24.4$.

Observando a figura 2.12, podemos ver duas trajetórias de evolução a partir de dois pontos muito próximos.

Inicialmente as duas trajetórias parecem coincidentes mas, a partir de certa altura, a divergência é grande. Por causa dos seus termos não-lineares, uma pequena variação na localização na condição inicial afecta enormemente as trajetórias obtidas, uma característica do comportamento matemático a que se chama caos.

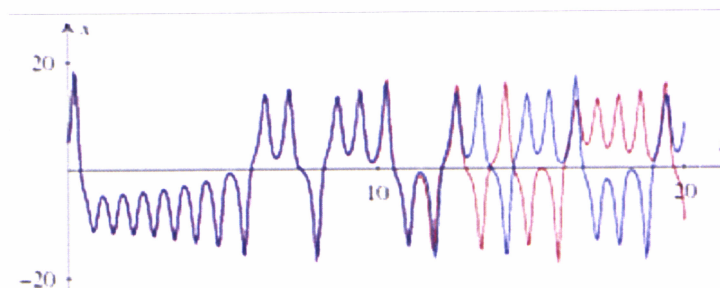


Figura 2.12: Duas trajetórias de evolução a partir de dois pontos que diferem um do outro em apenas 10^{-5} na coordenada x .

Num sistema dinâmico, o caos surge quando dois pontos iniciais arbitrariamente perto um do outro divergem exponencialmente, de tal modo que o seu comportamento futuro é eventualmente imprevisível.

Lorenz descobriu que, para certos valores dos parâmetros das equações, o sistema nunca tende para um comportamento previsível a longo prazo e que, por essa razão, não é possível também fazer previsões do tempo meteorológico a longo prazo.

As suas descobertas ganharam notoriedade quando publicou um artigo denominado *Predictability: Does a flap of butterfly's wing in Brazil set off a tornado in Texas?*, onde ficou célebre o conceito de efeito Borboleta. Se o bater das minúsculas asas de uma borboleta no Brasil podia desencadear, semanas depois, um tornado no Texas, isso deve-se à extrema sensibilidade às condições iniciais de um sistema tão complexo como o clima. Trata-se de um sistema caótico e a mais ínfima variação nas condições iniciais pode produzir comportamentos a longo prazo muito diferentes.

Para $r < 1$, todas as trajetórias, independentemente das condições iniciais, aproximam-se eventualmente do ponto fixo na origem. Usando uma linguagem introduzida acima, dizemos que todo o espaço xyz é a bacia de atracção para o atractor, neste caso, na origem.

Para valores $1 < r < 13.93$, o sistema tem três pontos fixos.

O sistema tem um dos pontos fixos na origem ($x = 0, y = 0, z = 0$). De facto, o sistema não se move a partir desse ponto. Mas basta que uma das coordenadas seja ligeiramente diferente de 0 para que o sistema evolua para um dos outros pontos fixos. Ou seja, a origem é aquilo a que se chama um ponto fixo repulsor. Os outros dois pontos fixos são atractivos se r não for muito grande. O comportamento do sistema é estável e tende para um de dois pontos fixos. Para $r = 1$ temos um ponto de bifurcação.

Quando $r < 24.06$ não temos atractor caótico, mas se $r > 13.926$ existem órbitas caóticas produzindo caos transiente.

Se o parâmetro r é maior do que 24.74, o comportamento do sistema muda radicalmente e os pontos fixos perdem a sua estabilidade. Deixam de ser atractores para serem repulsores mas a trajectória nunca se pode afastar muito da zona da figura anterior. A trajectória vai de um lado para o outro sendo repelida pelos três pontos fixos de modo que acaba por se dobrar sobre si mesma de um modo muito complexo e sem nunca se intersectar a si mesma.

Um ponto do atractor de Lorenz nunca repete o seu movimento, traça continuamente anéis em forma de asas de borboleta. O movimento do ponto é caótico pois não é possível prever em qual dos anéis e em que local destes ele se irá encontrar. Por exemplo, pode traçar um anel na asa esquerda, depois saltar para a direita e traçar aí dois anéis, voltar para a esquerda e de salto em salto vai percorrendo o atractor.

É ao conjunto de pontos em que a trajectória se encontra que se chama o atractor que neste caso recebe o nome de *atractor estranho*.

Neste modelo a duplicação do período aumenta enquanto o valor do parâmetro r

diminui.

A construção deste novo conjunto de ideias continuou com os contributos de cientistas brilhantes como Benoit Mandelbrot ou Mitchell Feigenbaum, que descobriram coisas tão inesperadas como a universalidade e a auto-semelhança em sistemas caóticos.

A universalidade resultou dos trabalhos de Feigenbaum, que detectou inesperadamente a mesma constante de proporcionalidade na evolução de diferentes sistemas não-lineares.

A auto-semelhança está intimamente ligada com o desenvolvimento da geometria fractal, que representa o lado mais conhecido e belo do Caos.

Como é típico dos atractores caóticos, o atractor de Lorenz exhibe auto-semelhança e é um objecto fractal. (Se introduzíssemos no programa a possibilidade de seleccionarmos uma zona limitada do atractor e fizéssemos zooms sucessivos, vizualizando pormenores com uma resolução cada vez maior, veríamos que o aspecto da curva seria sempre semelhante). Embora muitos fractais não sejam caóticos, como o triângulo de Serpinsky ou a curva de Koch, muitos fenómenos caóticos exibem estruturas fractais nos seus atractores estranhos ou nas sucessivas bifurcações de transição para o caos, como na função logística (2.2).

2.2 Caracterização dos sistemas caóticos

Muitas vezes dizemos que os sistemas dinâmicos são caóticos sem uma definição precisa do que isso significa. Nesta secção vamos abordar conceitos relacionados com a natureza caótica de aplicações e fornecer algumas definições de conjuntos invariantes caóticos.

Definição 2.2.1 *Uma função $f : D \rightarrow D$ diz-se topologicamente transitiva em D se para qualquer par de conjuntos abertos U, V que intersectam D existe um z em $U \cap D$ e um número natural n tal que $f^n(z)$ está em V . Equivalentemente, f é topologicamente*

transitiva em D se quaisquer dois pontos x e y em D e qualquer $\varepsilon > 0$, existe um z em D tal que $|z - x| < \varepsilon$ e $|f^n(z) - y| < \varepsilon$ para algum n .

Intuitivamente, uma aplicação transitiva topologicamente tem pontos nos quais eventualmente após iterações movem-se de um aberto arbitrariamente pequeno para qualquer outro.

Se iterarmos um ponto do intervalo $]0, 1[$, através da função logística (2.2) com $a = 4$, há órbitas que parecem percorrer todos os pontos possíveis do intervalo. O mesmo não sucedia com os outros valores de a .

Uma forma intuitiva de interpretar transitividade topológica é subdividir o intervalo em subintervalos e exigir que, iterando, seja possível atingir qualquer subintervalo, partindo de qualquer subintervalo. Consideremos a função logística (2.2) com $a = 4$ e a partição do intervalo I em 10 subintervalos: $I_k = [\frac{k-1}{10}, \frac{k}{10}]$. Começando no intervalo I_2 , após 4 iterações já tínhamos visitado todos os intervalos. Informalmente, topologicamente transitiva é quando podemos chegar a qualquer lado a partir de qualquer lado.

Proposição 2.2.1 *Seja $f : D \rightarrow D$. Se f é topologicamente transitiva em D , então ou D é um conjunto infinito ou D consiste numa órbita periódica.*

Demonstração: Suponhamos que D contém um número finito de pontos e $f : D \rightarrow D$ é topologicamente transitiva. Escolhendo qualquer ponto em D e designá-lo por x_0 . Vamos verificar que x_0 é periódico e qualquer ponto em D está na órbita de x_0 . Seja $\varepsilon > 0$ tal que a distância entre quaisquer dois pontos em D seja pelo menos ε . Se z e y estão em D e $z \neq y$, então $|z - y| \geq \varepsilon$. Tal ε existe desde D seja finito. Seja y um elemento de D . Como f é topologicamente transitiva, existe um z em D e um número natural m tal

que $|x_0 - z| < \varepsilon$ e $|f^m(z) - y| < \varepsilon$. Mas ε é a distância mínima entre pontos distintos em D logo $|z_0 - z| < \varepsilon$ implica que $z = x_0$ e $|f^m(z) - y| < \varepsilon$ implica que $f^m(z) = y$. Estes argumentos juntos implicam que $f^m(x_0) = f^m(z) = y$. Agora, escolhamos y_0 um ponto em D outro que não x_0 . Por um argumento semelhante ao procedimento anterior, podemos mostrar que existe um n tal que $f^n(y_0) = x_0$. Como já sabemos que existe um m tal que $f^m(x_0) = y_0$, torna-se claro que, $f^{n+m}(x_0) = f^n(y_0) = x_0$, e logo x_0 é periódico com período $n + m$. Por contraposição, se f não consiste na órbita do ponto periódico e f é topologicamente transitiva em D , então D é um conjunto infinito.

Uma das características fundamentais dos sistemas caóticos, é a sensibilidade às condições iniciais, ou seja, a possibilidade de estados tão próximos quanto se queira, sofrerem evoluções muito diferentes.

Definição 2.2.2 *Um sistema dinâmico $f : D \rightarrow D$ diz-se exibir sensibilidade às condições iniciais num ponto $x \in D$ se existe $r > 0$ tal que, para cada $\varepsilon > 0$ existe $y \in D$ e $k \geq 0$ tais que $d(x, y) < \varepsilon$ e $d(f^k(x), f^k(y)) \geq r$, onde $d(\cdot)$ é uma distância definida em D . Se f exibe sensibilidade às condições iniciais em todos os pontos de D , então diz-se que exibe sensibilidade às condições iniciais.*

Isto é, se temos dois sistemas iguais que iniciam nas condições iniciais x e $x + \varepsilon$, respectivamente, onde ε é uma quantidade muito pequena, as suas órbitas divergem rapidamente uma da outra, crescendo a sua separação exponencialmente.

A figura 2.13 mostra a evolução de duas órbitas para o mesmo período de tempo cujas condições iniciais diferem apenas 10^{-3} para o valor de x . Inicialmente as trajectórias parecem coincidir, no entanto, a partir de um certo tempo, a diferença entre as duas órbitas

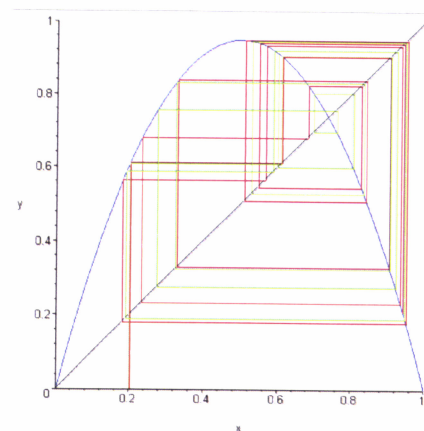


Figura 2.13: Evolução de duas órbitas que diferem apenas de 10^{-3} .

crece rapidamente. Este é também um sistema que revela sensibilidade às condições iniciais.

Um aspecto curioso da sensibilidade às condições iniciais é que este tipo de comportamento das aplicações era já conhecido, mas, porque sempre num contexto onde esse comportamento era exactamente algo a evitar, nunca suscitou muita curiosidade por parte dos matemáticos desse tempo. Já em 1908, Poincaré reflectiu sobre esse fenómeno, como é referido na sua obra *Science et Méthode*:

" Uma causa muito pequena, que nos escapa, determina um efeito considerável que não conseguimos ver, e então dizemos que esse efeito é devido ao acaso. Se conhecessemos exactamente as leis da natureza e a situação do universo no instante inicial, poderíamos prever com exactidão a situação desse mesmo universo num instante posterior. Mas, mesmo que as leis naturais já não tivessem segredos para nós, só aproximadamente poderíamos conhecer a situação inicial. Se isso nos permite prever a situação posterior com o mesmo grau de aproximação, e isto nos basta, dizemos que o fenómeno foi previsto, que é regido pelas

leis; mas isso nem sempre acontece; pequenas diferenças nas condições iniciais podem gerar diferenças muito grandes nos fenómenos finais; um pequeno erro nas primeiras produziria um erro enorme nos segundos. A previsão torna-se impossível."

No entanto, foi apenas na década de 60, que se ouviu falar de sistemas sensivelmente dependentes das condições iniciais, através de Edward Lorenz.

Existem diferentes definições de sistemas caóticos. Vamos considerar a definição introduzida em [13], por Devaney.

Definição 2.2.3 *A função $f : D \rightarrow D$ é caótica se:*

- a) os pontos periódicos de f são densos em D .*
- b) f é topologicamente transitiva.*
- c) f exibe sensibilidade às condições iniciais.*

Teorema 2.2.1 *Seja D um subconjunto infinito de números reais e $f : D \rightarrow D$ contínua. Se f é topologicamente transitiva em D e os pontos periódicos de f são densos em D , então f é caótica em D .*

Demonstração: É suficiente provar que f exibe sensibilidade às condições iniciais. Começamos por mostrar que existe $\delta > 0$ tal que para cada x em D existe um ponto periódico q tal que $|x - f^i(q)| \geq 4\delta$ para todo o i . Note que um subconjunto denso de um conjunto infinito deve ser infinito. Como os pontos periódicos de f são densos em D , isto implica que há infinitos pontos periódicos, escolhemos dois, q_1 e q_2 pontos periódicos que não são da mesma órbita. Seja δ o maior número positivo satisfazendo $\delta \leq \frac{1}{8} |f^n(q_1) - f^m(q_2)|$ para todo o m e n .

Claramente, $\delta > 0$, pois q_1 e q_2 são periódicos e há apenas um número finito de pontos da forma $f^n(q_1)$ e $f^m(q_2)$. Logo, há apenas um número finito de distâncias da forma $|f^n(q_1) - f^m(q_2)|$, e assim δ é um oitavo da distância mais curta. Agora se x é um ponto qualquer em D , então a desigualdade triangular implica que

$$\delta \leq \frac{1}{8} |f^n(q_1) - f^m(q_2)| \leq \frac{1}{8} (|f^n(q_1) - x| + |x - f^m(q_2)|)$$

ou

$$8\delta \leq |f^n(q_1) - x| + |x - f^m(q_2)|$$

para todos os n, m . Mas isto implica que pelo menos um $|x - f^n(q_1)| \geq 4\delta$ para todo o n ou $|x - f^m(q_2)| \geq 4\delta$ para todo o m . Logo, existe um ponto periódico q tal que $|x - f^i(q)| \geq 4\delta$ para todo o i . De seguida, escolhendo $\delta > 0$, para cada ponto em D existe um ponto periódico que se move através da iteração de f pelo menos a uma distância mínima 4δ . Seja x um ponto arbitrário em D e ε tal que $0 < \varepsilon < \delta$. Fixando um ponto periódico q satisfazendo $|x - f^i(q)| \geq 4\delta$ para todo o i . Em seguida, usamos q para provar que existe um w em D satisfazendo $|w - x| < \varepsilon$ e $|f^n(x) - f^n(w)| > \delta$ para algum n . Isto implica que f tenha sensibilidade às condições iniciais.

Exemplo 2.2.1 *Consideremos o exemplo simples de uma função linear:*

$$f(x) = cx, c > 1$$

Dado um valor inicial positivo, depois de n iteradas teremos

$$x_n = c^n x_0$$

Se adicionarmos um erro E_0 a x_0 , temos $u_0 = x_0 + E_0$ obtemos $u_n = c^n(x_0 + E_0)$ e o erro ao fim de n passos é

$$E_n = u_n - x_n = c^n(x_0 + E_0) - c^n x_0 = c^n E_0$$

O erro inicial é aumentado por um factor c em cada iterada, o sistema é sensível, mas não é certamente caótico.

Dado um desvio inicial num sistema caótico, este tornar-se-á tão grande como o verdadeiro sinal, isto é, depois de um certo número de iteradas, o erro terá a mesma ordem de grandeza que os valores correctos.

Para o iterador linear do exemplo (2.2.1), a relação entre o sinal, os valores da iteração e o erro é sempre a mesma. O erro relativo é

$$\frac{E_n}{x_n} = \frac{c^n E_0}{c^n x_0} = \frac{E_0}{x_0} = \text{constante.}$$

Isto indica que a iteração é bem comportada e que o erro é perfeitamente controlado, apesar do sistema apresentar sensibilidade às condições iniciais.

A sensibilidade por si só não implica caos. No entanto, esta propriedade é uma qualidade que todos os sistemas caóticos têm e detectá-la é já um passo grande na compreensão do caos. Outros sistemas poderão não ter esta qualidade. Como qualquer outra grandeza, também a sensibilidade precisa de ser quantificada.

Para deduzir uma medida para a velocidade de crescimento do erro por iterada temos a relação

$$\left| \frac{E_n}{E_0} \right| = c^n \Leftrightarrow \ln \left(\left| \frac{E_n}{E_0} \right| \right) = n \ln(c) \Leftrightarrow \ln(c) = \frac{1}{n} \ln \left| \frac{E_n}{E_0} \right|$$

donde se retira o valor de c ,

$$c = e^{\frac{1}{n} \ln \left| \frac{E_n}{E_0} \right|}.$$

Fez-se um estudo em [[38], pg 515] onde se observou que os valores deste expoente para a função logística com $a = 4$ tomando diferentes valores iniciais andava à volta de 0,7 o que dá $c \approx e^{0,7} \approx 2$, ou seja, para a função logística 2.2, para $a = 4$, em cada iteração, um pequeno erro aproximadamente duplica. Evidentemente, isto só se aplica para erros suficientemente pequenos e só acontecem em média. Intervalos perto do vértice da parábola são comprimidos, ao passo que intervalos perto dos pontos 0 e 1 aumentam cerca de 4 vezes. Este raciocínio leva-nos directamente para o conceito de expoente de Lyapunov $\lambda(x_0)$. Este quantifica o crescimento médio de erros infinitesimais na condição inicial x_0 , que voltaremos a abordar na secção seguinte.

Vamos assumir que E_0 pode ser arbitrariamente pequeno. Como se verifica a igualdade seguinte podemos tentar estimar cada factor em separado.

$$\frac{E_n}{E_0} = \frac{E_n}{E_{n-1}} \cdot \frac{E_{n-1}}{E_{n-2}} \cdots \frac{E_1}{E_0}$$

Supondo que tal é possível, então para calcular o valor desta expressão teríamos de multiplicar todos os factores individuais, o que causaria um grande erro em computadores normais. Felizmente é possível evitar esta situação visto que

$$\frac{1}{n} \ln \left| \frac{E_n}{E_0} \right| = \frac{1}{n} \ln \left(\frac{E_n}{E_{n-1}} \cdot \frac{E_{n-1}}{E_{n-2}} \cdots \frac{E_1}{E_0} \right) = \frac{1}{n} \sum_{k=1}^n \ln \left| \frac{E_k}{E_{k-1}} \right|$$

Somando os logaritmos evitamos problemas maiores nos arredondamentos. Trataremos agora de estimar cada parcela. Consideremos $\left| \frac{E_{k+1}}{E_k} \right|$. Este número descreve o quanto um pequeno erro E_k em x_k é aumentado (ou reduzido) na iteração seguinte.

Este factor de ampliação do erro é, na sua essência, independente do tamanho do erro E_k .

Se considerarmos, por exemplo, um erro em x_k que é metade de E_k , então é expectável que o erro em x_{k+1} seja também metade do que devia, $\frac{E_{k+1}}{2}$. Para justificar este raciocínio consideremos um erro $E_k = \varepsilon$ em x_k , isto é, $\widetilde{x}_k = x_k + \varepsilon$.

Voltemos à função logística definida em (2.2) para $a = 4$.

$$f(x) = 4x(1 - x)$$

então, o erro E_{k+1} em x_{k+1} é

$$E_{k+1} = f(\widetilde{x}_k) - f(x_k) = \widetilde{x}_{k+1} - x_{k+1} = 4(x_k + \varepsilon)(1 - x_k - \varepsilon) - 4x_k(1 - x_k) = 4\varepsilon(1 - 2x_k) - 4\varepsilon^2.$$

Dividindo por $E_k = \varepsilon$ obtemos

$$\frac{E_{k+1}}{E_k} = 4(1 - 2x_k) - 4\varepsilon,$$

uma quantidade dominada pelo primeiro termo, dado que 4ε é comparativamente pequeno.

Por isso se vê que o factor de ampliação do erro no passo $k + 1$ é independente do erro $E_k = \varepsilon$ desde que ε seja pequeno. Esta independência fornece-nos uma solução para o cálculo que queremos efectuar.

Fixamos um erro arbitrariamente pequeno ε e estimamos os factores de ampliação $\frac{E_{k+1}}{E_k}$ para erros E_0 muito pequenos, estimando $\frac{|E_{k+1}|}{\varepsilon}$ onde

$$E_{k+1} = f(x_k + \varepsilon) - f(x_k) \text{ e } f(x) = 4x(1 - x),$$

Em suma, temos uma fórmula melhorada para determinar o expoente de Lyapunov, que abordaremos na secção seguinte, baseado num número indeterminado, tão grande quanto se queira, de iterações.

A fórmula é

$$\frac{1}{n} \ln \left| \frac{E_n}{E_0} \right| \approx \frac{1}{n} \sum_{k=1}^n \ln \left| \frac{\widetilde{E}_k}{\varepsilon} \right|.$$

Como $\left| \frac{\widetilde{E}_k}{\varepsilon} \right| = \frac{|f(x_{k-1+\varepsilon}) - f(x_{k-1})|}{\varepsilon}$. Se f for diferenciável da expressão anterior, passando ao limite, obtemos

$$\lim_{E_0 \rightarrow 0} \frac{1}{n} \ln \left| \frac{E_n}{E_0} \right| \approx \frac{1}{n} \sum_{k=1}^n \ln \left| f'(x_{k-1}) \right| \quad (2.4)$$

Dado um desvio inicial num sistema caótico, este tornar-se-á tão grande como o verdadeiro sinal, isto é, depois de um certo número de iteradas, o erro terá a mesma ordem de grandeza que os valores correctos.

É do nosso conhecimento que existem pontos que têm um comportamento de natureza completamente diferente: os pontos periódicos.

Se começarmos a iteração num desses pontos visitaremos apenas um número restrito de intervalos, apesar de infinitas vezes.

Teoricamente para a função logística (2.2) é possível encontrar infinitos pontos deste tipo em cada intervalo. Além da sensibilidade e transitividade topológica, a existência de pontos periódicos em cada subintervalo é uma condição necessária para a existência de caos. Para $a = 4$ temos, por exemplo, a órbita de período 3.

$$\left\{ \sin^2 \frac{\pi}{7}, \sin^2 \frac{2\pi}{7}, \sin^2 \frac{4\pi}{7} \right\}$$

Contudo, se tomarmos um destes 3 valores para início da iteração, e se tomarmos uma precisão simples dos números no computador, o erro é aproximadamente 0.00000005 (dependendo do processador) e o erro, em média, duplica em cada iterada. Isto dá um erro de aproximadamente $1/20$ em 20 iteradas.

Observamos que há um desvio visível da órbita periódica a partir da iterada 30. A sensibilidade do sistema não permite detectar ciclos instáveis.

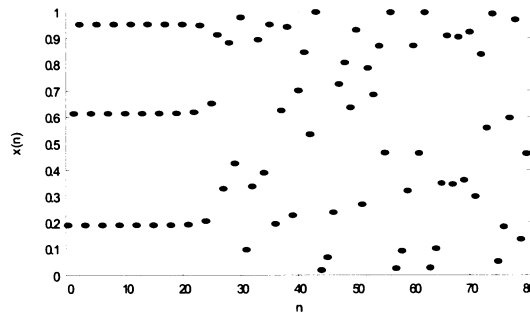


Figura 2.14: Série temporal para $a = 4$ e $x_0 = \sin^2(\pi/7)$.

Continuamos o nosso estudo com a função logística (2.2) mostrando que esta função é caótica em $[0, 1]$. Infelizmente, analisar através da definição é uma tarefa difícil. De seguida, vamos mostrar que as dinâmicas desta função (2.2) em $[0, 1]$ são as mesmas que as dinâmicas da função tenda em $[0, 1]$. Dizemos, neste caso, que a função logística é topologicamente conjugada com a função tenda em $[0, 1]$.

Definição 2.2.4 *Sejam $f : D \rightarrow D$ e $g : E \rightarrow E$ funções. Então f é topologicamente conjugada com a função g se existe um homeomorfismo $\tau : D \rightarrow E$ tal que $\tau \circ f = g \circ \tau$. Neste caso, τ é chamada de conjugação topológica.*

Teorema 2.2.2 *Sejam D e E subconjuntos de números reais, $f : D \rightarrow D$, $g : E \rightarrow E$, e $\tau : D \rightarrow E$ uma conjugação topológica de f e g . Logo*

- a) $\tau^{-1} : E \rightarrow D$ é uma conjugação topológica;
- b) $\tau \circ f^n = g^n \circ \tau$ para todos os números naturais n ;
- c) p é ponto periódico de f se e só se $\tau(p)$ é um ponto periódico de g . Mais, os períodos primitivos de p e $\tau(p)$ são idênticos;
- d) Se p é ponto periódico de f com $W^s(p)$ onde $W^s(p)$ é o conjunto estável de p , então o conjunto estável de $\tau(p)$ é $\tau(W^s(p))$.

e) Os pontos periódicos de f são densos em D se e só se os pontos periódicos de g são densos em E .

f) f é topologicamente transitiva em D se e só se g é topologicamente transitiva em E .

g) f é caótico em D se e só se g é caótico em E .

Podemos usar a conjugação topológica para mostrar que a função logística é caótica em $[0, 1]$.

Teorema 2.2.3 A função logística $f_4 : [0, 1] \rightarrow [0, 1]$ definida em (2.2) para $a = 4$ é topologicamente conjugada com a função tenda $T : [0, 1] \rightarrow [0, 1]$ definida por:

$$T(x) = \begin{cases} 2x & \text{para } x \text{ em } [0, \frac{1}{2}] \\ 2 - 2x & \text{para } x \text{ em } [\frac{1}{2}, 1] \end{cases}$$

Demonstração: É suficiente mostrar que a aplicação $\tau(x) = \sin^2\left(\frac{\pi}{2}x\right)$ é uma conjugação topológica entre as duas funções. Em particular, devemos mostrar que:

a) $\tau : [0, 1] \rightarrow [0, 1]$ é injectiva e sobrejectiva

b) $\tau : [0, 1] \rightarrow [0, 1]$ é contínua

c) $\tau^{-1} : [0, 1] \rightarrow [0, 1]$ é contínua, e

d) $\tau \circ T = f \circ \tau$

As três primeiras propriedades seguem do facto que $\tau'(x)$ existe em $[0, 1]$ e é estritamente positivo em $[0, 1]$. E a última propriedade segue directamente das identidades trigonométricas elementares.

No livro [25] demonstra-se que relativamente à função T , os pontos periódicos de T são densos em $[0, 1]$, T é topologicamente transitiva em $[0, 1]$, $T : [0, 1] \rightarrow [0, 1]$ exhibe sensibilidade às condições iniciais e que a função T é caótica em $[0, 1]$.

Corolário 2.2.1 A função logística (2.2) é caótica em $[0, 1]$.

A demonstração é uma consequência imediata do teorema (2.2.3) e, do teorema (2.2.2) e do facto da função T ser caótica em $[0, 1]$.

2.3 Grandezas que caracterizam os sistemas dinâmicos

Como podemos ter a certeza de que o que parece caótico realmente o é, que não é apenas muito complicado, mas ainda assim, perfeitamente previsível?

Ao olharmos para um conjunto de dados experimentais aparentemente caótico, como podemos ter a certeza de que este não é periódico, só que com um período extremamente longo? Ou seja, como será possível identificar os sinais de caoticidade dos dados? E poderá o caos ser medido? Porquê que o queremos quantificar?

De seguida, vamos apresentar diferentes maneiras de responder quantitativamente a estas questões.

As razões que nos levam a quantificar comportamento caótico é que nos permite distinguir comportamento caótico de comportamento não caótico mas com ruído. Estas grandezas ajudam-nos a determinar quantas variáveis são necessárias para modelar o sistema. Mudanças nestas grandezas podem indicar mudanças do comportamento dinâmico do sistema.

2.3.1 Expoentes de Lyapunov

A teoria dos Sistemas Dinâmicos tem as suas raízes no século vinte. E não apenas Poincaré, sem dúvida o mais importante deles todos, mas também outros houve que contribuíram também com conceitos e resultados que mais tarde viriam a integrar o corpo principal da Teoria dos Sistemas Dinâmicos. Um desses matemáticos foi Aleksandr Mikhailovich Lyapunov (1857 – 1918).

Matemático e engenheiro russo, Lyapunov foi aluno de P.L. Chebyshev e colega de um outro grande matemático, A.A. Markov. Entre as suas contribuições para a Matemática, de uma forma geral destacam-se trabalhos sobre equações diferenciais, estabilidade de sistemas e teoria das probabilidades, mas, podemos dizer que o seu trabalho fundamental se concentrou na estabilidade do equilíbrio e do movimento de um sistema mecânico. Na sua tese de doutoramento, o problema geral da estabilidade do movimento, assumiu uma grande importância para o estudo da dinâmica não linear, limites de estabilidade e estabilidade de movimentos, e é nesse sentido que introduz o conceito que lhe permite medir a forma como duas trajectórias, inicialmente muito próximas, se afastam, ou se aproximam, uma da outra no futuro. Essa ideia, hoje justamente associada ao seu nome, refere-se aos expoentes de Lyapunov.

Um método usado para determinar se um sistema é caótico é estimar o expoente de Lyapunov.

Isto dá-nos uma indicação da convergência ou divergência de duas órbitas inicialmente muito próximas.

Para um sistema de dimensão n , verifica-se que o número de expoentes de Lyapunov é n , que no seu conjunto formam o chamado *Espectro de Lyapunov*.

A natureza qualitativa do atractor é representada por um padrão de sinais das componentes do espectro de Lyapunov, o qual indica a taxa na qual trajectórias próximas de atractores mudam ao longo do tempo em cada uma das diferentes direcções ortogonais. O 0 indica que não houve mudança, + e - indicam divergência e convergência, respectivamente.

Para atractores de dimensão 3, estes padrões são representados por 3 elementos, vector

$(-, -, -)$ para ponto fixo, vector $(0, -, -)$ para ciclo-limite, $(0, 0, -)$ para torus e $(+, 0, -)$ para caos.

Para atratores de dimensão 4, há diferentes tipos de caos, $(+, 0, -, -)$ para caos, $(+, +, 0, -)$ para hipercaos e $(+, 0, 0, -)$ para 2-torus caótico.

Pelo menos uma componente positiva fornece-nos necessariamente, mas não suficiente, evidência qualitativa para caos.

Os expoentes de Lyapunov fornecem uma caracterização qualitativa e quantitativa do comportamento caótico de sistemas não lineares.

Lyapunov para aplicações unidimensionais

Supomos que temos duas cópias do modelo logístico (2.2), em uma usamos a variável x e na outra y , sendo idênticas excepto que as suas condições iniciais serem diferentes. Começamos com x_0 e y_0 muito próximos.

Quando iteramos o sistema a partir de dois valores iniciais x_0 e y_0 os valores começam a distanciar-se no início lentamente e depois mais rapidamente. E eventualmente mostram que não há nenhuma co-relação entre elas, embora as dinâmicas dos dois surjam da mesma equação.

Como vimos atrás, este "alongamento para além" da distância entre pontos vizinhos iniciais é chamado *sensibilidade às condições iniciais*. Uma maneira de caracterizar o sistema dinâmico caótico é medir o comprimento desta sensibilidade dependente.

Consideremos que temos dinâmicas determinísticas perfeitas, $x_{t+1} = f(x_t)$, com $x_t \in \mathbb{R}$ e duas condições iniciais x_0 e y_0 cuja separação $|x_0 - y_0|$ seja muito pequena, a separação

depois da primeira iteração é

$$|x_1 - y_1| = |f(x_0) - f(y_0)| = \frac{|f(x_0) - f(y_0)|}{|x_0 - y_0|} |x_0 - y_0| \approx |f'(x_0)| |x_0 - y_0|$$

fazendo uso da definição $f'(x_0) = \lim_{y_0 \rightarrow x_0} \frac{f(x_0) - f(y_0)}{x_0 - y_0}$.

O expoente de Lyapunov local λ^* em x_0 é definido de modo que $e^{\lambda^*} = \left| \frac{x_1 - y_1}{x_0 - y_0} \right|$, ou ainda $\lambda^* \approx \ln \left| \frac{x_1 - y_1}{x_0 - y_0} \right| \approx \ln |f'(x_0)|$.

A quantidade $\left| \frac{x_1 - y_1}{x_0 - y_0} \right|$ chamada *número de Lyapunov local* mede a expansão das órbitas a partir de x_0 . O valor absoluto assegura que o seu logaritmo seja um número real. Se $\frac{x_1 - y_1}{x_0 - y_0}$ for negativo, significa que depois da iteração, dois pontos próximos trocam a sua ordem, ou seja, o maior passa a menor e vice-versa.

O facto de se saber como os números de Lyapunov variam no espaço, permite-nos identificar regiões do atractor com muita ou pouca previsibilidade para pequenos erros iniciais.

No entanto, sendo a órbita o objecto mais importante no estudo da dinâmica de um sistema, é mais interessante termos uma ideia global dos valores dos números de Lyapunov que os sucessivos pontos dessa órbita vão assumindo. Esse valor, associado a uma órbita do sistema, é designado de Expoente de Lyapunov.

Como queremos o número que descreve a sensibilidade de uma função como um todo, e não apenas da primeira iteração. Constatamos que

$$|x_2 - y_2| \approx |f'(x_1)| |x_1 - y_1| \approx |f'(x_1)| |f'(x_0)| |x_0 - y_0|$$

e por iteração,

$$|x_n - y_n| \approx \left| \prod_{t=0}^{n-1} f'(x_t) \right| |x_0 - y_0|$$

Note que a solução da equação às diferenças linear $x_{t+1} = ax_t$ é $x_n = a^n x_0$ e que o cálculo da separação por iteração (no sistema linear é a) é $\left(\left| \prod_{t=0}^{n-1} f'(x_t) \right| \right)^{\frac{1}{n}}$. Isto é a média geométrica das quantidades $|f'(x_t)|$. O termo do *expoente de Lyapunov* é o logaritmo deste cálculo da separação das iterações.

Logo, o procedimento para quantificar a sensibilidade às condições iniciais de uma função às diferenças de dimensão 1 é o seguinte:

1. Iteramos a função para gerar a sequência de valores de x_0, x_1, \dots, x_{n-1} .
2. Calculamos o declive da função em cada um dos pontos x_0, x_1, \dots, x_{n-1} .
3. Calculamos o valor absoluto da média geométrica dos valores do passo 2.

Este valor representa a sensibilidade às condições iniciais de uma função como um todo.

Finalmente, se usarmos a regra do logaritmo para as derivadas de n iteradas e tomando o limite quando n tende para infinito obtemos:

$$\lambda = \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=0}^{n-1} \ln |f'(x_i)| \quad (2.5)$$

O factor de expansão média e^λ , chama-se número de Lyapunov global. Estes têm propriedades diferentes dos expoentes de Lyapunov locais. O expoente de Lyapunov global determina a taxa média exponencial de separações de duas órbitas inicialmente próximas, ou ainda, a expansão média do espaço de fases.

Se o expoente de Lyapunov de uma aplicação unidimensional é positivo implica a existência de caos. Se o expoente de Lyapunov é negativo, então existe um ponto fixo ou uma órbita periódica atractiva. E se o expoente de Lyapunov é zero, ocorre bifurcação no comportamento do sistema.

O interesse fundamental deste trabalho consiste no cálculo de expoentes de Lyapunov para conjunto de dados, obviamente finitos. Assim sendo, desde logo se pode colocar a questão da convergência (não da sua existência) do limite acima apresentado para sequências finitas e não muito grandes da órbita de uma condição inicial.

Exemplo 2.3.1 *Consideremos a função logística (2.2), definida no intervalo $0 \leq x \leq 1$ e com $a = 4$. Para esta aplicação é possível determinar analiticamente o expoente de Lyapunov, ver [41], obtendo-se $\lambda = \ln 2 \approx 0.693147$. Vejamos então que valor se obtém escolhendo um qualquer valor inicial, por exemplo $x_0 = 0.04$ e calculando aproximações para λ considerando um número finito N , cada vez maior, mas sempre não muito grande, de parcelas em (2.5). Observamos a convergência de três casas decimais para o valor de $\lambda \approx 0.693$.*

Neste âmbito, é também muito interessante calcular o expoente de Lyapunov para a família de aplicações logísticas em função do seu parâmetro.

Recuperando o comentário anterior, relativamente ao significado do sinal do expoente de Lyapunov, é muito curioso sobrepôr o gráfico da variação do expoente de Lyapunov com o Diagrama de Bifurcação. Deste modo, é possível observar as janelas de estabilidade e os pontos de bifurcação. No gráfico do lado esquerdo, observamos os Expoentes de Lyapunov para a função logística $f_a(x_n) = ax_n(1 - x_n)$ em função do seu parâmetro a com $2.8 < a < 4$. No gráfico do lado direito observamos o Diagrama de Bifurcação da mesma função para os mesmos parâmetros a .

A tabela (2.6) mostra os resultados dos expoentes de Lyapunov calculados na função logística (2.2) para os diferentes valores do parâmetro a .

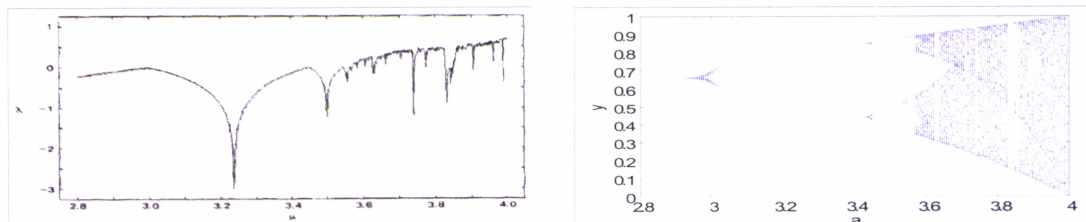


Figura 2.15: Expoentes de Lyapunov e o diagrama de bifurcação da função logística.

a	λ
0.5	-0.69317
1	-0.00034
2.1	-2.30248
2.5	$-\ln 2 \approx -0.69316$
3	-0.00023
3.5	-0.87200
3.8	0,26075
4	$\ln 2 \approx 0.69316$

(2.6)

Os resultados tabela (2.6) foram calculados com o programa MAPLE com 50000 iterações e arredondados a 5 casas decimais.

De facto, observamos que à medida que o sistema se torna mais complexo, maior é o valor do expoente de Lyapunov.

Expoentes de Lyapunov para aplicações bidimensionais

Os conceitos de Lyapunov podem ser alargados para aplicações em \mathbb{R}^n . Para a generalização da definição formal dos expoentes de Lyapunov para um sistema num espaço de dimensão n arbitrária teremos naturalmente que considerar um disco infinitesimal de dimensão n , com centro num certo ponto X_0 , e a matriz das primeiras derivadas da aplicação f em causa, $Df(X_n)$, calculada no ponto $X_n = f^n(X_0)$. Temos portanto os objectos matemáticos necessários para escrevermos a definição pretendida. Mas, antes

disso, vejamos o significado da generalização das ideias na origem do conceito de Exponentes de Lyapunov.

Para tal, consideremos um sistema descrito por uma aplicação no plano.

Consideremos o sistema dinâmico $f : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ e um qualquer ponto $X_0 \in \mathbb{R}^2$. A questão que se coloca agora é a de saber qual o comportamento, perante a iteração por f , dos pontos pertencentes a um disco infinitesimal com centro em X_0 e raio ε . Como facilmente se observa, o disco de pontos iniciais surge, após a iteração dos seus pontos pela aplicação f , deformado dando origem uma elipse, ou seja, cujos eixos maior e menor marcam as direcções de expansão e contracção, respectivamente, do sistema, e a partir dos quais qualquer outro comportamento entre estes pode ser analisado. Deste modo, aquilo que vai corresponder ao conceito inicial do expoente de Lyapunov, uma medida de afastamento ou aproximação de órbitas, num instante infinitesimalmente próximas, serão as taxas médias destes números de expansão e contracção medidos pelos eixos principais de cada uma destas elipses infinitesimais obtidas pela deformação de discos infinitesimais pela iteração da aplicação.

Como a área da elipse de eixos a e b é dada por $A = \pi ab$, então $A_0 = \pi\varepsilon^2$, se o raio do disco inicial for igual a ε , após a 1ª iteração temos $A_1 = \pi\varepsilon e^{\lambda_1}\varepsilon e^{\lambda_2}$, então temos que $\frac{A_1}{A_0} = e^{\lambda_1 + \lambda_2}$, onde vem que $\lambda_1 + \lambda_2 = \ln\left(\frac{A_1}{A_0}\right)$. O valor absoluto do determinante da matriz jacobiana, o jacobiano, representa o factor de expansão da área, assim temos que, $\lambda_1 + \lambda_2 = \ln|\det(J_f)|$, onde J_f representa a matriz jacobiana da aplicação f . A quantidade $\det(J_f)$ geralmente varia ao longo da órbita, pelo que, para determinarmos a soma dos expoentes de Lyapunov globais teremos de fazer a média da equação anterior, ou seja, $\lambda_1 + \lambda_2 = \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=0}^{N-1} \ln|\det J_f(X_n)|$.

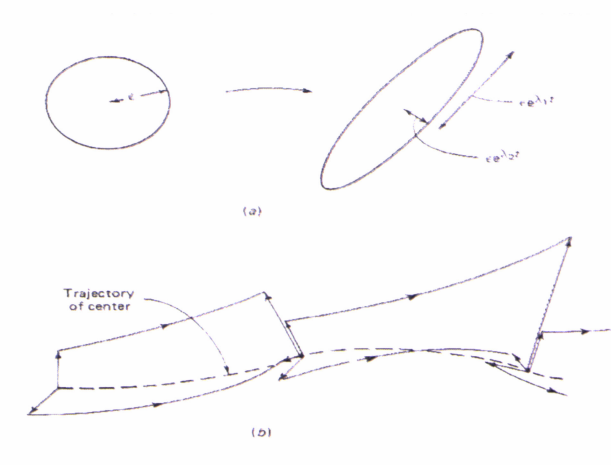


Figura 2.16: Divergência de órbitas para um sistema bidimensional.

Ora, recordando como os valores próprios da matriz J_f medem exactamente os comprimentos dos eixos maior e menor da elipse descrita anteriormente, temos naturalmente que, neste caso, dois expoentes de Lyapunov, são dados por $\lambda_j = \lim_{N \rightarrow \infty} \frac{1}{N} \ln |\Lambda_j^N|$, $j = 1, 2$ onde Λ_j^N representa o j -ésimo valor próprio da matriz $J_f(X_i)$, sendo $J_f(X_i)$ a matriz jacobiana de f calculada no ponto X_i .

Desde que a noção de atractores estranhos tem sido introduzida, tem sido claro que os expoentes de Lyapunov podem ser usados no seu estudo. Consideremos uma bola infinitesimalmente pequena no espaço de fases F -dimensional. Durante a evolução pode começar a alterar, mas sendo infinitesimal, pode ficar um elipsóide. Designemos o eixo principal desta elipsóide por $\epsilon_i(t)$ ($i = 1, \dots, F$). Os expoente de Lyapunov λ_i são então determinados por

$$\epsilon_i(t) \approx \epsilon_i(0) e^{\lambda_i t}. \quad (2.7)$$

A soma dos λ_i , que descrevem a contracção de volume, tem de ser negativo. Mas visto que o atractor estranho resulta de um processo de encolhimento e de alongamento, requer que pelo menos um dos λ_i seja positivo. Inversamente, um expoente de Lyapunov positivo implica sensibilidade às condições iniciais e por conseguinte ter um comportamento caótico.

O maior expoente de Lyapunov

Um teste muito usual para detectar caos num sistema é o cálculo do maior expoente de Lyapunov. Se o expoente é positivo indica a presença de caos no sistema. Quando temos as equações que geram o sistema é fácil o seu cálculo. A ideia principal é seguir duas órbitas muito próximas e calcular a taxa média de separação das duas.

Na figura 2.17 podemos observar o seguinte:

Quando as órbitas se afastam demasiado, uma das órbitas tem de ser arrastada para a vizinhança da outra, ao longo da linha de separação.

O procedimento é efectuado em cada iteração. O procedimento completo é o seguinte:

1. Iniciar com uma condição inicial que pertence à bacia de atracção.

O ideal seria escolhermos uma condição inicial que estivesse no atractor. Neste caso, o passo 2 seria omitido.

2. Iterar até que a órbita pertença ao atractor.

Para alguns sistemas é suficiente iterar algumas centenas de vezes. O erro causado pelo facto de estar ligeiramente fora do atractor não é, geralmente, grande a não ser que esteja muito próximo de um ponto de bifurcação.

Este passo requer algum conhecimento prévio sobre o sistema em estudo.

3. Seleccionar um ponto muito próximo, separado por uma distância de d_0 .
4. Iterar uma vez mais, e calcular a nova separação d_1 .

A separação é calculada pela soma dos quadrados das diferenças para cada variável.

Logo para um sistema de dimensão 2 com variáveis x e y , a separação deve ser

$$d = [(x_a - x_b)^2 + (y_a - y_b)^2]^{\frac{1}{2}},$$

onde os índices a e b representam as duas órbitas, respectivamente.

5. Determinar $\ln |d_1/d_0|$, ou em qualquer outra base do logaritmo mais conveniente.

6. Reajustar uma das órbitas de modo que a separação seja d_0 na mesma direcção que d_1 .

Por exemplo, num sistema de dimensão dois, supondo que a órbita que têm de ser ajustada o seu valor após uma iteração é (x_{b1}, y_{b1}) .

Terá de ser reinicializado em

$$x_{b0} = x_{a1} + d_0(x_{b1} - x_{a1})/d_1 \quad e \quad y_{b0} = y_{a1} + d_0(y_{b1} - y_{a1})/d_1.$$



Figura 2.17: Método numérico para o cálculo do maior expoente de Lyapunov.

7. Repetir os passos de 4 a 6 um grande número de vezes, calculando a média no passo

5.

É aconselhável desprezar os primeiros valores para assegurar que as órbitas sejam orientadas ao longo da direcção da expansão máxima.

2.3.2 Entropia

A entropia é uma grandeza que descreve a quantidade de ordem do sistema. Podemos dizer que quanto maior for a entropia do sistema, mais desorganizado ele está, e isso pode ser uma característica do comportamento caótico da série.

A entropia de um sistema dinâmico dá-nos uma ideia da complexidade desse sistema inerente ao grau de incerteza e à quantidade de informação envolvida.

A sequência de n lançamentos de uma moeda ao ar (de uma moeda equilibrada) tem 2^N resultados possíveis. Há uma incerteza nos resultados que vamos medir como entropia, que é o logaritmo do número de resultados possíveis.

$$S = \log 2^N = N \log 2 \quad (2.8)$$

Se tivermos que mandar uma mensagem por cada sequência possível temos de enviar 2^N mensagens.

A "capacidade de informação" deste exemplo é outra vez medido como o logaritmo do número de possíveis mensagens enviadas.

$$I = \log 2^N = N \log 2 \quad (2.9)$$

Neste contexto se usarmos a base 2 para o logaritmo dizemos então que temos N bits de informação.

Uma visão alternativa é que a medida de um resultado particular, isto é, sequências de caras e coroas diz-nos algo sobre o sistema e temos $N \log 2$ bits de informação.

2.3.3 Entropia nos Sistemas Dinâmicos

Podemos generalizar estas ideias a sistemas dinâmicos com N resultados possíveis.

A aparência complexa de várias representações gráficas de comportamentos caóticos leva-nos naturalmente à questão da relação entre estatística e caos. Uma maneira de relacionar este aspecto é aplicar o conceito de Entropia no sistema caótico comparando o resultado com o sistema estatístico associado.

Há diferentes maneiras de caracterizar comportamento caótico baseado em noções que são (pelo menos formalmente) relacionadas com o conceito de entropia.

Vamos ver que esta medida de entropia pode ser generalizada dando-nos um formalismo poderoso para caracterizarmos os sistemas dinâmicos.

Segundo [23], em primeiro lugar, vamos necessitar de contar estados. Vamos começar por dividir o espaço de estados do sistema em células, usualmente do mesmo tamanho. Para um sistema dissipativo (para o qual existe um atractor) precisamos de dividir apenas a região contendo o atractor. Para um sistema conservativo dividimos o intervalo percorrido pelas trajectórias.

Começamos com uma família de condições iniciais, usualmente todas localizadas dentro de uma célula. À medida que o sistema evolui no tempo, as trajectórias irão geralmente passar num grande número de células no espaço de estados. Após n unidades de tempo (cada um com duração τ) calculamos a frequência relativa (probabilidade) p_i na qual o sistema visita cada uma das células.

E a entropia do sistema é definida como sendo:

$$S_n = -k \sum_{i=1}^n p_i \ln p_i \quad (2.10)$$

Note que p_i é a probabilidade de uma dada trajectória com início numa célula inicial, cair na i -ésima célula após n unidades de tempo. No que se segue, vamos supor $k = 1$ para simplificar a notação. No caso de $p_i = 0$, consideramos $p_i \ln p_i = 0$ na fórmula (2.10).

Esta quantidade deve ser interpretada como uma quantidade de desordem do sistema ou informação necessária para especificar o estado do sistema.

Relativamente ao significado de probabilidade, interessa-nos especificar o seguinte: se iniciamos as trajectórias com M pontos na célula inicial e, depois de n unidades de tempo, temos M_i pontos das trajectórias na i -ésima célula. Logo podemos definir probabilidade como $p_i = \frac{M_i}{M}$, que é a fracção do número de pontos das trajectórias que caem na i -ésima célula naquele tempo. Antes de introduzirmos mais pormenores destas ideias, vejamos alguns casos simples para vermos como a definição de entropia se comporta.

Primeiro, supomos que todas as trajectórias se movem juntas de célula a célula à medida que o tempo evolui. Logo $p_i = 1$ para a célula ocupada e $p_i = 0$ para as células que não foram ocupadas. Neste caso

$$S_n = 0$$

para todo o n . Logo, entropia constante corresponde àquilo que interpretamos como desenvolvimento regular (digamos ao longo de um ciclo-limite).

Em segundo lugar, supomos que o número N_n de células ocupadas cresce com o tempo e que todas as células têm a mesma probabilidade, nomeadamente, $\frac{1}{N_n}$. Logo a entropia é dada por

$$S_n = \ln N_n$$

Vemos que a entropia cresce se o logaritmo do número de células ocupadas cresce.

Para um sistema aleatório, cada um dos M pontos da trajectória caem para a sua célula. Neste caso,

$$S_n = \ln M.$$

O importante é que este número cresce sem limite enquanto M se torna muito grande.

Por outro lado, para um sistema determinístico caótico ou regular com um número pequeno de variáveis (isto é, sistemas que não são completamente aleatórios) a entropia torna-se independente de M , para um M grande.

Estamos concentrados com as mudanças na entropia e não no seu valor. Por exemplo, se escolhermos um conjunto de condições iniciais em duas células e o seu desenvolvimento é regular, o valor da entropia não seria 0 mas permaneceria constante enquanto o sistema evolui. A mudança de entropia é conhecida pela entropia de Kolmogorov Sinai, às vezes, chamada $K-S$ entropia ou k -entropia para simplificar, a qual descreve a taxa de mudança da entropia enquanto o sistema evolui.

Definição 2.3.1 A k -entropia após n unidades de tempo é $K_n = \frac{1}{\tau}(S_{n+1} - S_n)$, K_n é a mudança da entropia desde $t = n\tau$ a $t = (n + 1)\tau$.

O que queremos actualmente é a k -entropia ao longo do atractor inteiro, em ordem a caracterizar o atractor como um todo. Logo, definimos a quantidade média K ,

$$K = \lim_{N \rightarrow \infty} \frac{1}{N\tau} \sum_{n=0}^{N-1} (S_{n+1} - S_n) = \lim_{N \rightarrow \infty} \frac{1}{N\tau} (S_N - S_0)$$

Ao considerarmos um N grande permitimos que as trajectórias evoluam um tempo longo, cobrindo todo o atractor. Vamos introduzir mais dois limites: Um dos limites faz o tamanho da célula tender para 0, ou seja, usamos divisões muito pequenas do espaço de estados. Este procedimento deve tornar K independente de como se divide o espaço. O outro limite faz tender o intervalo de tempo τ para 0, ou seja, usamos incrementos de tempo mais pequenos e logo obtemos uma descrição mais fina da dinâmica. Colocando todos estes limites juntos temos uma definição completa de k -entropia,

$$K = -\lim_{\tau \rightarrow 0} \lim_{l \rightarrow 0} \lim_{N \rightarrow \infty} \frac{1}{N\tau} (S_N - S_0)$$

Quando a k -entropia é aplicada a secções de Poincaré e iteramos as funções colocamos $\tau = 1$ e largamos o limite $\tau \rightarrow 0$.

Vamos ver outro exemplo, suponhamos que o número de células ocupadas N_n cresce exponencialmente ao longo do tempo. $N_n = N_0 e^{\lambda n \tau}$ e todas as células ocupadas têm a mesma probabilidade $p_r = \frac{1}{N_n}$. A k -entropia é igual a λ , o parâmetro que caracteriza o crescimento exponencial do número de células ocupadas. O número de células ocupadas, no entanto, é proporcional à distância entre as trajectórias que inicialmente ocuparam a mesma célula. Logo, vemos que λ é a média do expoente de Lyapunov para o sistema. Para o crescimento exponencial do número de células ocupadas e para células ocupadas com igual probabilidade, a k -entropia e o expoente de Lyapunov são o mesmo. Quando temos mais do que um expoente, a k -entropia torna-se igual à soma dos expoentes de Lyapunov positivos, o resultado é conhecido como Identidade de Pesin.

A figura 2.18 mostra os resultados de aplicar o conceito de entropia na função logística para vários valores do parâmetro.

Como esperado, a entropia no geral cresce com os diferentes valores de a , excepto quando há picos para baixo referentes a comportamentos periódicos.

A entropia de Kolmogorov (Kolmogorov, 1959) é uma medida importante através da qual podemos caracterizar o desenvolvimento caótico (numa dimensão arbitrária) no espaço de fases.

De seguida, podemos ver nas figuras 2.19 à 2.21 as K -entropias para trajectórias regulares, caóticas e aleatórias (dimensão um), segundo [44].

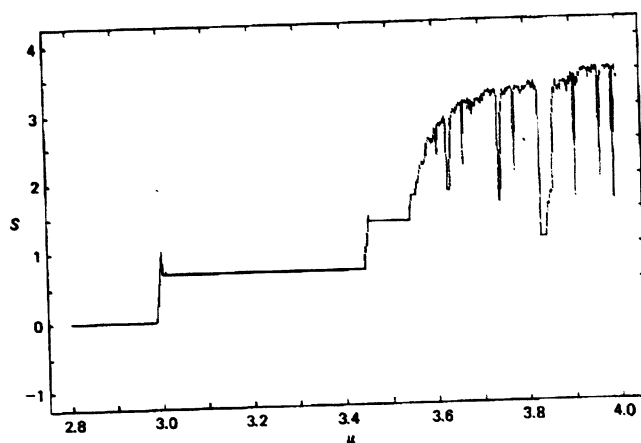


Figura 2.18: Entropia do sistema da função logística para os vários parâmetros de a . A entropia máxima correspondente a igual probabilidade para cada 40 células é 3.6888.

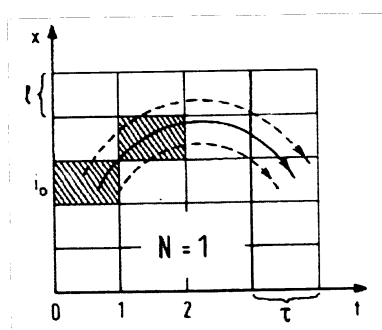


Figura 2.19: Trajetória Regular. Pontos inicialmente adjacentes mantêm-se adjacentes.

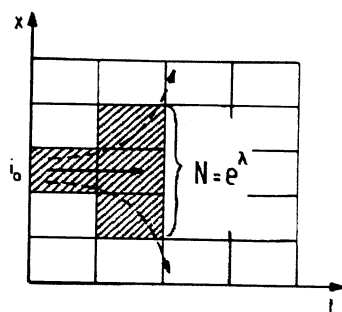


Figura 2.20: Trajetória caótica. Pontos inicialmente adjacentes tornam-se separados exponencialmente.

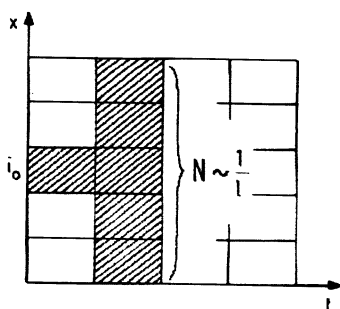


Figura 2.21: Trajetória aleatória. Pontos inicialmente adjacentes são distribuídos com igual probabilidade.

Definição alternativa de k-entropia Nas dinâmicas não lineares, é usada uma definição ligeiramente diferente da k -entropia. Nesta definição alternativa, tomamos uma trajectória simples ao longo de um grande período de tempo. Podemos cobrir a região do atractor do espaço de estados com células. De seguida, começamos a trajectória numa das células e chamamos-lhe célula $b(0)$. Passado τ tempo, o ponto da trajectória estará na célula $b(1)$. Em $t = 2\tau$, a trajectória estará na célula $b(2)$, e assim, sucessivamente até $t = N\tau$, registando uma sequência particular de células: $b(0), b(1), \dots, b(N)$.

E logo começamos com uma segunda trajectória a partir da mesma célula inicial. Porque geralmente obtemos sequências diferentes de células para a segunda sequência. Repetindo este processo muitas vezes, estamos a gerar um grande número de sequências. De seguida, vamos calcular o número relativo de vezes que uma sequência particular ocorre nas N células. Vamos denominar este número relativo por $p(i)$ para a i -ésima sequência.

Definição 2.3.2 Definimos a entropia como sendo: $S_N = \langle \sum_i p(i) \ln p(i) \rangle$.

Onde o somatório é tomado em todas as sequências de N células que iniciam na célula $b(0)$. Os sinais \langle e \rangle significam que o cálculo do somatório abrange todas as células no atractor.

Definição 2.3.3 Definimos a k -entropia como a taxa média de crescimento da entropia relativamente ao comprimento da sequência: $K = \lim_{N \rightarrow \infty} \frac{1}{N} (S_N - S_0)$.

Vamos observar alguns casos particulares. Em primeiro lugar, vamos assumir que todas as sequências iniciam na mesma célula inicial; neste caso temos um desenvolvimento regular: Todas as trajectórias partem da mesma célula e andam juntas ao longo do tempo. Logo, $S_N = 0$ para todo N e $K = 0$ desde que S não mude. No outro extremo, vamos

assumir que cada uma das M sequências ocorrem uma única vez (isto é, o sistema é puramente aleatório). Neste caso, $p(i) = \frac{1}{M}$ e $S_N = \ln M$ e cresce sem limite enquanto M cresce. Logo, vemos que a k -entropia cresce enquanto o número de trajectórias cresce. Vamos agora assumir que o número de trajectórias distintas nas células M_{seq} cresce exponencialmente (em média) com o comprimento N das sequências: enquanto as sequências se tornam maiores, elas afastam-se devendo-se a comportamento das trajectórias ser distinto com condições iniciais (ligeiramente) diferentes. Em termos mais formais, assumimos que $M_{\text{seq}} = e^{\lambda N}$. Obviamente, λ é o expoente de Lyapunov do sistema. E se assumirmos que cada trajectória distinta ocorre com a mesma probabilidade $p(j) = \frac{1}{M_{\text{seq}}}$, então facilmente vemos que a k -entropia é apenas $k = \lambda$. Logo, uma vez mais temos uma igualdade entre k -entropia e o expoente de Lyapunov positivo.

Relação de K com os Expoentes de Lyapunov Como já foi referido atrás, para correspondências de dimensão um, K é apenas o expoente de Lyapunov positivo. Em sistemas com dimensões mais altas, perdemos informação sobre o sistema porque a célula na qual estava previamente o sistema espalha-se para novas células no espaço de fases a uma taxa determinada pelos expoentes de Lyapunov positivos. Então temos que a taxa K à qual se perde informação do sistema é igual à soma dos expoentes de Lyapunov positivos (Pesin, 1977).

Quando os expoentes de Lyapunov são positivos, a divergência das órbitas ao longo do tempo vai gerando cada vez mais informação sobre o exacto valor da condição inicial. Portanto a entropia de Kolmogorov-Sinai mede o valor médio da taxa de produção de informação por unidade de tempo.

Em conclusão, o maior expoente de Lyapunov proporciona uma definição rigorosa de caos e a soma dos expoentes de Lyapunov positivos (ou equivalentemente a entropia de Kolmogorov-Sinai) caracteriza a complexidade dinâmica das órbitas do sistema.

Capítulo 3

Séries temporais

3.1 Introdução

Até agora vimos como diferentes comportamentos podem ser gerados nos sistemas dinâmicos, tais como pontos fixos, ciclos-limites e caos.

O nosso objectivo de aplicar as dinâmicas é relacionar os sistemas matemáticos a sistemas psicológicos de interesse.

O movimento caótico tem muitas propriedades definidas e por isso pode ser examinado usando várias técnicas. O nosso objectivo é encontrar métodos que nos permita diferenciar séries temporais com ruído aleatório das caóticas.

Antes de se usar alguma técnica para analisar a série temporal devemos inspeccionar visualmente a série e reparar se há algum facto relevante presente na série. No entanto, a inspecção visual tem desvantagens, como por exemplo, nem sempre é possível inspeccionar uma longa série temporal e determinar se a série se repete ou não.

De seguida, começaremos com uma sequência de medidas, isto é, uma série temporal e queremos observar as suas dinâmicas, por isso vamos introduzir algumas ferramentas de análise de séries temporais, seguiremos [1].

3.1.1 Reconstrução do atractor

As ideias sobre a reconstrução do espaço de fases através do método do atraso temporal (time delay) devem-se a Ruelle, Packard e Takens. Takens demonstrou que é possível recriar uma aplicação topologicamente equivalente ao comportamento de um sistema dinâmico multidimensional através de uma sucessão cronológica de uma única variável observada. Através desta técnica de reconstrução, são preservados os expoentes de Lyapunov e a dimensão do atractor.

Segundo [1], para expandir um sinal unidimensional para um espaço m -dimensional, mergulhamos o atractor num espaço de dimensão m , e as coordenadas de dimensão m , X_i , do atractor são construídos a partir de uma série temporal, x_i ($i = 1, 2, 3, \dots, N$), como se segue:

$$X_i = (x_i, x_{i+\xi}, x_{i+2\xi}, \dots, x_{i+(m-1)\xi}), \quad i = 1, 2, 3, \dots, N - (m-1)\xi$$

onde X_i é um vector de dimensão m no atractor produzido por amostras discretas de uma série temporal x_i , i é o índice de tempo, m é a dimensão de imersão e ξ é um incremento temporal discreto, ou simplesmente um atraso temporal. Como resultado, obtém-se um conjunto de vectores

$$X_1, X_2, \dots, X_{N-(m-1)\xi},$$

onde N é o comprimento do sinal original.

O atraso temporal entre as variáveis reconstruídas é assim $\gamma = \xi dt$, onde dt é o intervalo de tempo entre o i -ésimo ponto e $(i+1)$ -ésimo ponto na série temporal.

Podemos observar nos gráficos *a*) e *b*) da figura 3.1, a série temporal e o atractor reconstruído, respectivamente. A gráfico *a*) mostra um ponto j na série temporal (isto é,

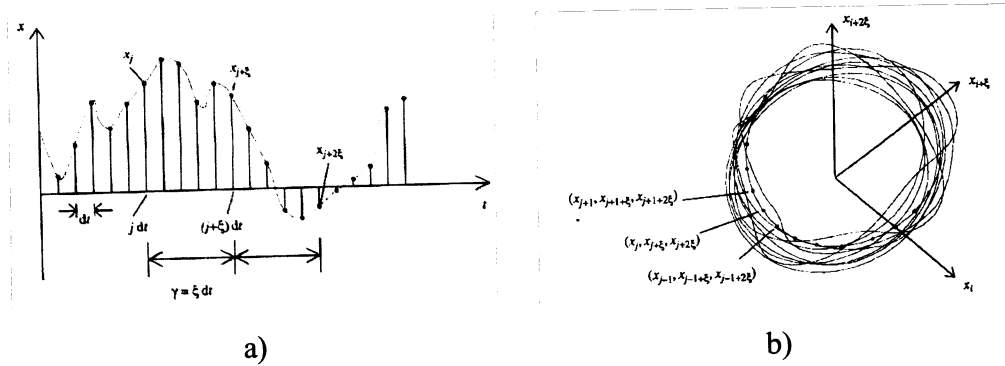


Figura 3.1: Reconstrução do atrator usando o método dos atrasos temporais. a) Série temporal em tempos discretos. b) Diagrama esquemático da reconstrução do atrator. Pode-se observar j -ésimo ponto e seus dois pontos vizinhos na trajetória.

$i = j$) e dois pontos com atraso temporal $j + \xi$ e $j + 2\xi$. Estes três pontos vão ser usados para reconstruir o j -ésimo ponto do atrator mergulhado num espaço de dimensão 3,

$X_j = (x_j, x_{j+\xi}, x_{j+2\xi})$. O ponto X_j pode ser observado no gráfico b) da figura 3.1 junto com dois pontos vizinhos X_{j-1} e $X_{j+\xi}$ na trajetória reconstruída.

Exemplo 3.1.1 Para reconstruir um atrator de dimensão 3 através da série temporal (1.01, 2.32, 1.93, 4.53, 2.22, 3.33, 3.34, 5.43, 5.55, 3.33, 4.02, ...) com um atraso temporal $\xi = 3$. Os primeiros 4 pontos do atrator reconstruído devem ser $X_i = (x_i, x_{i+\xi}, x_{i+2\xi})$, $i = 1, 2, 3, 4$, que são (1.01, 4.53, 3.34), (2.32, 2.22, 5.43), (1.93, 3.33, 5.55) e (4.53, 3.34, 3.33).

3.1.2 Cálculo da dimensão

Dimensão de "Box counting" A mais simples e de fácil compreensão é a definição de dimensão "Box counting", D_B . Para um atrator num hipervolume de dimensão um (comprimento)

$$D_B = \lim_{\delta \rightarrow 0} \left[\frac{\log(N)}{\log(1/\delta)} \right], \quad (3.1)$$

onde N é o número de hipercubos, com lados de comprimento δ , necessário para cobrir o atrator.

Nota: Somos livres de escolher qualquer base para os logaritmos. Em particular, optamos por logaritmos de base 10. A escolha da dimensão dos hipercubos é igual à dimensão do espaço de fases. Logo, para encontrar a dimensão "Box counting" no sistema de Lorenz (2.3) são usados hipercubos (caixas) de dimensão 3.

Dimensão de Informação Outro cálculo da dimensão citado frequentemente é a dimensão de informação, D_I . Tal como na dimensão de "box counting", o atrator é coberto com hipercubos de comprimento δ . Desta vez, porém, em vez de simplesmente contar cada cubo que contém parte do atrator, quantas vezes é que o atrator está contido em cada cubo. Esta medida destina-se a contar as diferenças na distribuição dos pontos cobrindo o atrator, é definida como:

$$D_I = \lim_{\delta \rightarrow 0} \left[\frac{I(\delta)}{\log(1/\delta)} \right], \quad (3.2)$$

onde, $I(\delta)$ é dada pela fórmula da entropia de Shannon, $I(\delta) = -\sum_i^N P_i \ln(P_i)$, onde P_i é a probabilidade de parte do atrator ocorrer dentro do i -ésimo hipercubo com lados de comprimento δ . Em particular, se no atrator a distribuição dos pontos tem igual probabilidade, $P_i = \frac{1}{N}$. Logo, $I(\delta) = \log(N)$, e neste caso, a equação (3.2) reduz-se à dimensão de "Box counting" da equação (3.1).

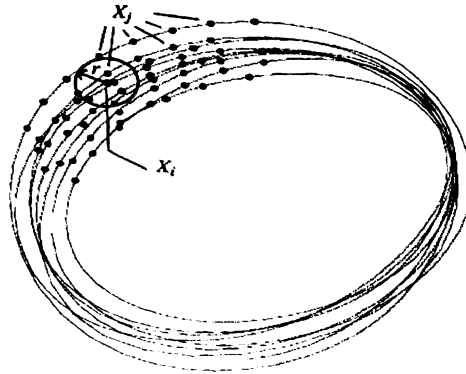


Figura 3.2: Amostra de uma hipersfera no atrator.

Logo, D_B conta simplesmente todos os hipercubos contendo partes do atrator e D_I solicita quantas vezes o atrator está dentro de cada hipercubo.

Dimensão de Correlação O cálculo de D_B , D_I , na prática, requer um tempo demasiado longo. A dimensão fractal mais usada para o cálculo da dimensão de atratores é a dimensão de correlação, D_c , que é mais eficiente computacionalmente e relativamente rápida quando implementada como um algoritmo para estimar a dimensão. Para definir D_c primeiro precisamos de definir o integral de correlação,

$$C_r = \frac{1}{N(N-1)} \sum_{i=1}^N \sum_{j=1; i \neq j}^N \theta(r - |X_i - X_j|) \quad (3.3)$$

onde θ é a função de Heaviside, r é o raio de uma hipersfera de dimensão n centrada em cada ponto da trajectória de um atrator, X_i , $i = 1, 2, \dots, N$. X_i é um vector multidimensional sendo o i -ésimo ponto do espaço de fases do atrator, isto é, para um atrator no espaço de fases de dimensão três com coordenadas x , y e z , $X_i = (x_i, y_i, z_i)$.

Podemos observar na figura 3.2 um esboço da definição 3.3, onde a hipersfera tem como centro um dos pontos da trajectória no atrator, X_i .

E contamos o número de outros pontos que estão dentro da hipersfera. É isto que faz a função de Heaviside na equação 3.3. A função de Heaviside é igual a 1 se o valor dentro dos parêntesis é positivo, e é 0 se o valor dentro dos parêntesis é negativo. X_i são os pontos da trajectória referida e X_j são outros pontos no atractor na vizinhança de X_i . $|X_i - X_j|$ é a distância de separação entre os dois pontos. $\theta(r - |X_i - X_j|)$ é igual a 1 se a distância entre dois pontos é menor que o raio da hipersfera; caso contrário é 0. Desta forma a função de Heaviside conta todos os pontos da hipersfera. O cálculo do integral de correlação da equação 3.3 envolve seguir a trajectória referida, parando em cada ponto discreto desta trajectória e contar o número de outros pontos atractores dentro da hipersfera de raio r . A soma cumulativa de todos os pontos obtidos é de seguida dividido por $N(N - 1)$, obtemos assim, o integral de correlação, C_r . O máximo valor de C_r é 1 e este máximo é obtido quando o raio da hipersfera é maior que o diâmetro maior do atractor e todos os pontos são contados. O valor mínimo de C_r é $\frac{2}{N(N-1)}$ e isto é encontrado quando contamos apenas dois pontos próximos no atractor. Note que C_r conta os pontos vizinhos próximos duas vezes.

Como vamos ver mais à frente o integral de correlação varia com o comprimento do raio da hipersfera com a lei de potência da forma:

$$C_r \sim r^{D_C}, \quad (3.4)$$

onde o expoente, D_C , é a dimensão de correlação. Ao aplicarmos este método para diferentes raios das hipersferas, D_C é obtido através do declive da região de escala no gráfico $\log(r)$ contra $\log(C_r)$. A *região de escala* é a região linear principal da curva de correlação. Em geral, uma boa estimativa de D_C requer uma grande quantidade de pontos, N .

A dimensão de correlação, dimensão de informação e dimensão de "box counting" estão relacionadas. De facto, fazem parte de uma grande colecção de dimensões, conhecidas como dimensões de Renyi. Pode-se mostrar que $D_C \leq D_I \leq D_B$. Pode-se afirmar que D_C e D_I aproximam-se assintoticamente do valor de D_B à medida que o atractor se torna mais distribuído uniformemente no espaço de fases. Se os pontos do atractor são distribuídos uniformemente então $D_C = D_I = D_B$.

Segundo [28], o comportamento caótico não tem período. Aperiodicidade significa que o mesmo estado nunca se repete duas vezes.

A presença de ruído numa série temporal de um sistema verdadeiramente periódico pode parecer aperiódica. Como os sistemas que não têm período podem diferir na sua não periodicidade, parece ter significado quantificar o quanto não é periódica.

Se um sistema não tem período, então as variáveis num estado nunca retornam aos seus valores prévios exactos. No entanto, há variáveis que talvez retornem para valores muito próximos dos anteriores.

Podemos perguntar "Quanto próximo?" e "Quantas vezes?". Já que muitas vezes não medimos todas as variáveis do estado de um sistema, muitas vezes não temos acesso a uma série temporal $\{X_n\}$ de vectores F -dimensionais.

Em lugar disso temos apenas um ou quanto muito alguns componentes de X_n . Isto é particularmente relevante para experiências reais onde o número de graus de liberdade é frequentemente muito alto se não infinito. Tais sistemas contudo podem ter atractores de baixa dimensão. É então muito desejável possuímos um método seguro que permita uma caracterização desses atractores para uma série temporal de variável única. $\{x_i, i = 1, \dots, N; x_i \in \mathbb{R}\}$, precisamos de usar a técnica de mergulho que representa todos os dados

medidos das variáveis de um estado.

De notar que D_t é a medida no tempo t .

Mergulhando a série temporal, criamos a sequência $\mathbf{D}_t = (D_t, D_{t-h}, \dots, D_{t-(p-1)h})$, onde p é a dimensão mergulhada e h é o atraso. Cada \mathbf{D}_t é um ponto de dimensão p do espaço e a série temporal pode ser considerada como uma sequência de pontos, um ponto em cada tempo t . Cada ponto representa o estado do sistema naquele instante.

Podemos calcular a distância entre dois pontos no tempo i e j : $\delta_{i,j} = |\mathbf{D}_i - \mathbf{D}_j|$.

Se a série temporal é periódica com período T então $\delta_{i,j} = 0$ quando $|i - j| = nT$ para $n = 0, 1, 2, \dots$

Em contraste, para séries temporais que não têm período, $\delta_{i,j}$ não mostra este padrão.

Suponhamos que escolhemos uma certa distância r , e queremos saber quando

$$|\mathbf{D}_i - \mathbf{D}_j| < r.$$

Uma maneira de fazer isto é construir um gráfico onde i está no eixo horizontal, j está no eixo vertical e um ponto é colocado na coordenada (i, j) se $|\mathbf{D}_i - \mathbf{D}_j| < r$.

Tais gráficos são chamados de *gráficos de recorrência* porque eles mostram como a trajectória reconstruída se move ou se repete.

Nas figuras a) e b) da figura 3.3 os pontos a preto são pontos tais que $|\mathbf{D}_i - \mathbf{D}_j| < r$. A trajectória das duas figuras tem período 4, logo o gráfico de recorrência consiste em segmentos na diagonal separados por 4. A figura b) tem a mesma trajectória que a figura a), mas agora temos o r dez vezes maior $r = 0,01$. O gráfico é idêntico ao anterior.

Para uma *série temporal periódica* de período T , o gráfico para um r muito pequeno tem um conjunto de segmentos de 45° , separadas por uma distância T na direcção vertical e horizontal.

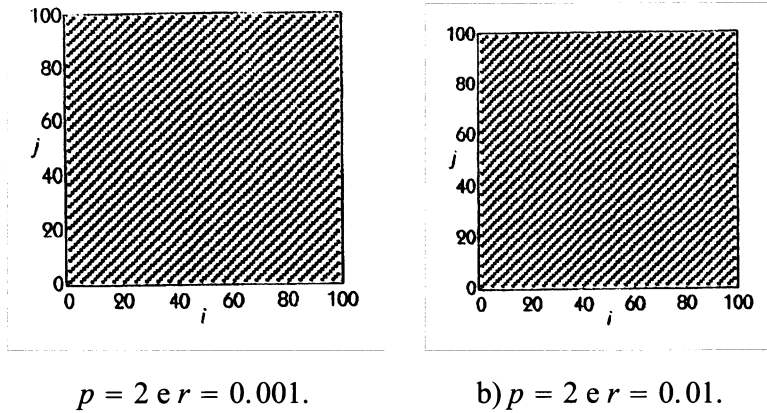


Figura 3.3: As figuras a) e b) são gráficos de recorrência da função quadrática $x_{t+1} = 3.52x_t(1 - x_t)$.

Para uma *série temporal caótica*, o gráfico de recorrência tem uma estrutura mais complicada, às vezes com trajetórias quase periódicas, há segmentos paralelos de 45° .

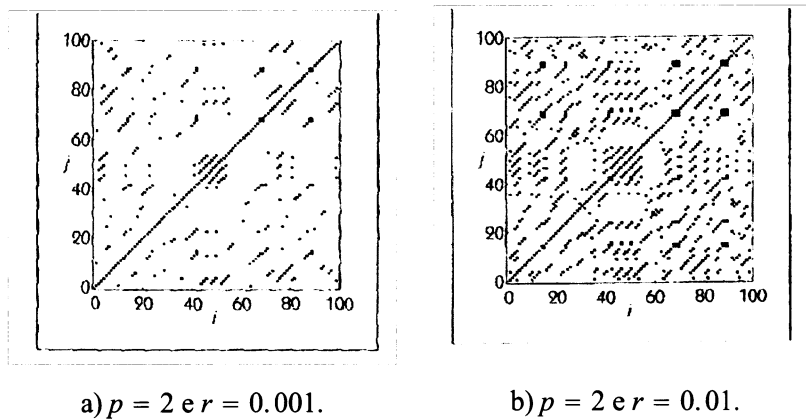


Figura 3.4: As figuras a) e b) são gráficos de recorrência para séries temporais gerados por $x_{t+1} = 4x_t(1 - x_t)$.

Relativamente à figura 3.4 a figura b) refere-se à mesma trajetória que a figura a), mas agora temos o r dez vezes maior $r = 0,01$.

Para números gerados aleatoriamente, tal estrutura não é evidente, ver figura 3.5

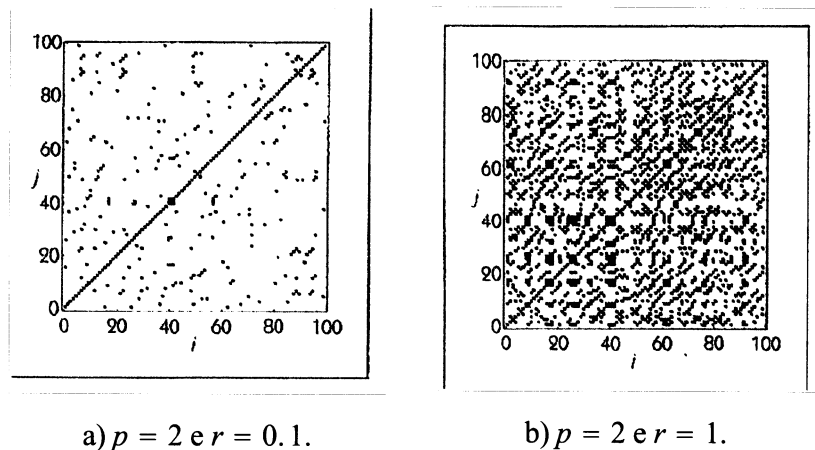


Figura 3.5: As figuras a) e b) são gráficos de recorrência para ruído-branco aleatório.

Novamente a figura b) corresponde à mesma trajetória que a a), mas com o r dez vezes maior.

De notar que o número de pontos num gráfico de recorrência diz-nos quantas vezes quaisquer dois pontos da trajetória têm uma distância inferior a r um do outro.

Podemos pensar em C_r como a densidade de tinta num gráfico de recorrência.

O integral de correlação é uma das quantidades mais importantes na análise de séries temporais caóticas.

O que é importante não é o valor de C_r em qualquer valor singular r , mas como C_r muda com r .

Quando r aumenta, mais pontos aparecem nos gráficos de recorrência e logo C_r aumenta.

Para um sistema periódico perfeito, aumentando pouco o r não muda o número de pontos significativamente.

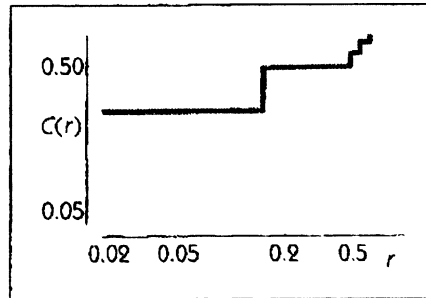


Figura 3.6: O integral de correlação C_r de série temporais periódicas com período 4 gerada por $x_{t+1} = 3.52x_t(1 - x_t)$, $N = 100$ são usados 100 pontos e a dimensão mergulhada é $p = 2$.

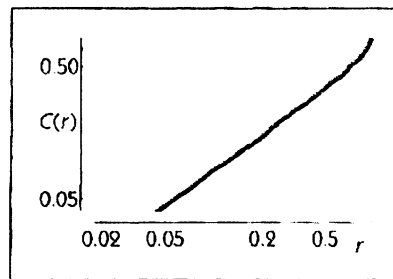


Figura 3.7: O integral de correlação C_r de série temporal caótica gerada por $x_{t+1} = 4x_t(1 - x_t)$, $N = 100$ são usados 100 pontos e a dimensão mergulhada é $p = 2$.

Para *dados caóticos* do modelo logístico (2.2), $x_{t+1} = ax_t(1 - x_t)$, aumentando r a mesma quantidade aparecem mais pontos, mas o crescimento é mais notório em dados com *ruído branco*.

Como podemos observar nas figuras 3.6 à 3.8, C_r é plana para o sistema periódico, tem um declive suave para o sistema caótico e tem um declive íngreme para um sistema aleatório.

De notar que C_r é desenhada numa escala logarítmica.

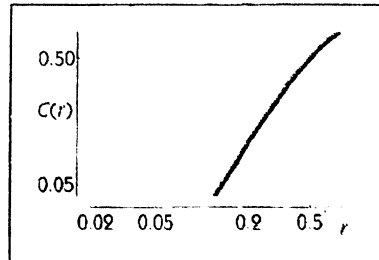


Figura 3.8: O integral de correlação $C(r)$ de série temporal produzidas por computador gerando números aleatórios, $N = 100$ e a dimensão mergulhada é $p = 2$.

Na figura 3.9 podemos observar o seguinte:

Se tivermos um conjunto de pontos dispersos mais ou menos uniformemente numa curva de dimensão um, se escolhermos um dos pontos como referência, quando r aumenta, o número de pontos dentro da distância r irá crescer linearmente com o comprimento de r .

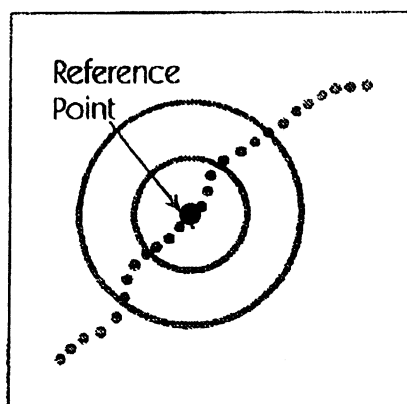


Figura 3.9: Pontos são dispersos ao longo de uma curva de dimensão 1.

Se tivermos um conjunto de pontos dispersos mais ou menos numa superfície de di-

mensão dois, escolhendo um dos pontos como referência, podemos ver que o número de pontos com uma distância menor que r do ponto de referência serão relacionados com a área do círculo de raio r , que é πr^2 .

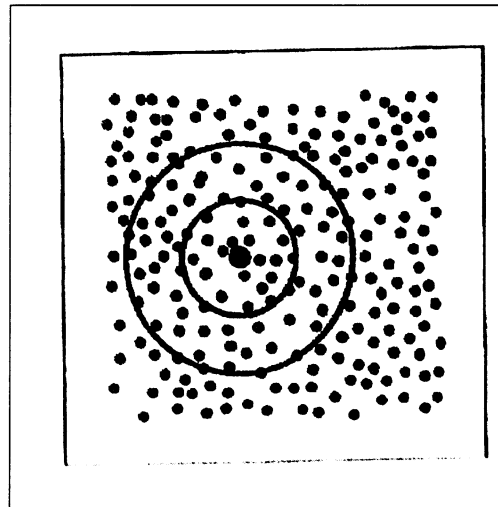


Figura 3.10: Os pontos estão dispersos ao longo de superfície de dimensão 2.

Podemos também observar na figura 3.10 o seguinte:

Se os pontos estão dispersos em todo o volume de dimensão 3, o número de pontos com uma distância menor que r de um ponto referência serão relacionados com o volume da esfera de raio r , que é $\frac{4}{3}\pi r^3$.

Em geral, para pontos dispersos em todo o objecto de dimensão D_C , o número de pontos mais próximos do ponto de referência a uma distância menor que r é proporcional a r^{D_C} .

Para calcular o integral de correlação para um conjunto de pontos, utilizamos cada um dos pontos de referência e contamos quantos outros estão dentro da distância r .

Isto sugere que o integral de correlação de pontos dispersos em toda a parte de um

volume de dimensão D_C será proporcional a r^{D_C} , que é $C_r = Ar^{D_C}$, onde A é a constante de proporcionalidade. Tomando o logaritmo dos dois lados temos $\log C_r = D_C \log r + \log A$.

Em ordem a encontrar D_C , simplesmente precisamos de colocar num gráfico $\log C_r$ em função de $\log r$ e encontrar o declive da linha resultante, como já foi referido atrás.

Este procedimento também pode ser aplicado para estimar a dimensão fractal de um objecto, em substituição da técnica do "box-counting".

Uma das aplicações da *dimensão de correlação* em séries temporais é o estudo de atractores em séries temporais.

A ideia inicial, proposta por [28], foi baseada na observação de que os atractores de sistemas caóticos são frequentemente auto-semelhantes e podem ser descritos através da sua dimensão fractal.

Um dos objectivos centrais é estabelecer que para pequenos r , C_r cresce como uma potência $C_r \sim r^{D_C}$ e que este "expoente de correlação" pode ser tomado como uma medida mais útil da estrutura local de um atrator estranho.

Dado um sinal experimental, se determinarmos $C_r \sim r^{D_C}$ com $D_C < F$, em que F é a dimensão do espaço de fases, sabemos que o sinal detém mais caos determinístico do que ruído aleatório, visto que o ruído aleatório resulta sempre em $C_r \sim r^F$.

Se a série temporal vem de uma órbita de um sistema dinâmico num atrator, a trajectória feita a partir da série temporal mergulhada terá as mesmas propriedades topológicas que o atrator original se a dimensão de mergulho for suficientemente grande. Em particular, a trajectória reconstruída terá a mesma dimensão que a original.

Funções de dimensão 1 Os casos mais simples de sistemas caóticos são representados por funções de um intervalo em si próprio, como por exemplo a função logística(2.2),

$$x_{n+1} = ax_n(1 - x_n)$$

Estudámos esta função quando $a = 3.5$, $a = 3.5699456$, $a = 3.6$, $a = 3.8282$ e para $a = 4.0$. No caso de $a = 3.5699456$...da função logística é apresentado na figura 3.11.

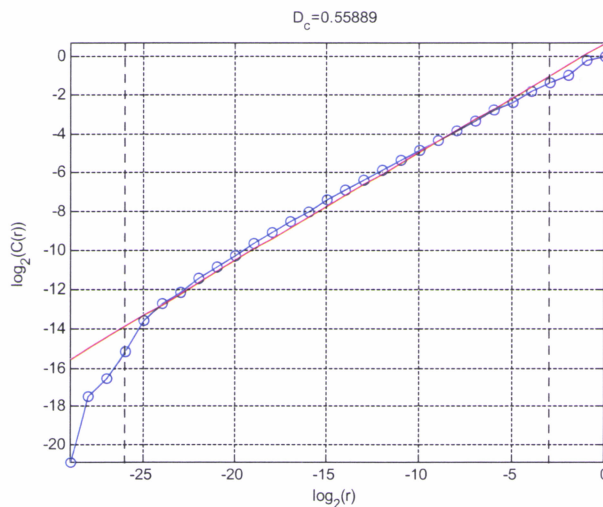


Figura 3.11: Dimensão de correlação para $a = 3.5699456$...

Da série temporal, gerada por um atrator, descartamos os primeiros 1000 pontos. Depois, com os restantes 1000, calculamos a distância entre todos os pares de pontos e contamos quantos pares de pontos estão a uma distância menor que r , sendo N o número de pontos considerados. Tomamos um N grande e aproximamo-nos do limite. Finalmente construímos o gráfico de $\log(C(r))$ em função de $\log(r)$, calcula-se, por interpolação, a melhor recta, cujo declive será aproximadamente D_C .

Para distâncias muito pequenas, os dados para $C(r)$ desviam-se de uma potência, mas isto era esperado: o comportamento para $a = 3.5699456$... não é caótico ainda, e por

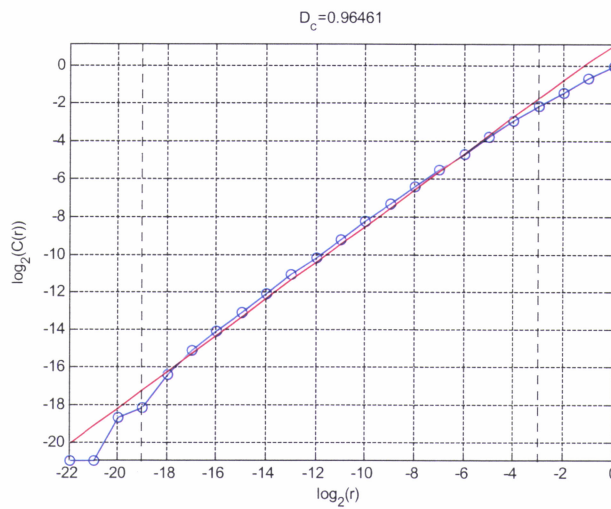


Figura 3.12: Dimensão de Correlação para $a = 4$.

consequente os valores x_n estão fortemente correlacionados. Verificamos que de facto a potência se mantém para valores pequenos de r se aumentarmos N ou usarmos somente valores $x_i, x_{i+p}, x_{i+2p}, x_{i+3p}, \dots$ com p sendo um número ímpar grande.

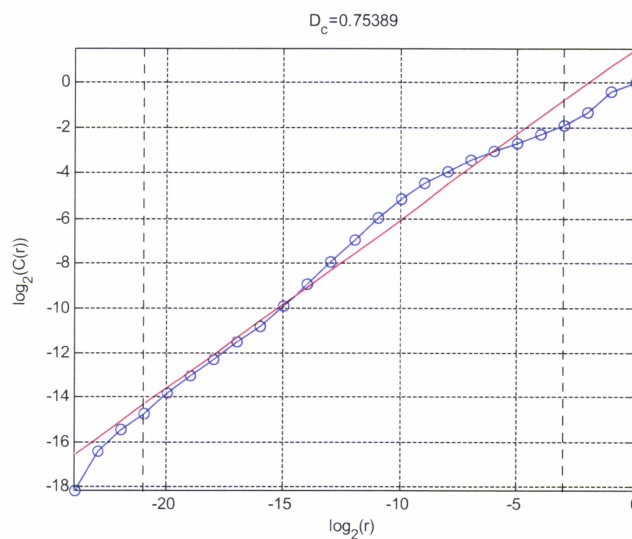


Figura 3.13: Dimensão de Correlação para $a = 3.8282$.

Vamos observar qual a dimensão de correlação de séries temporais para dados periódicos-

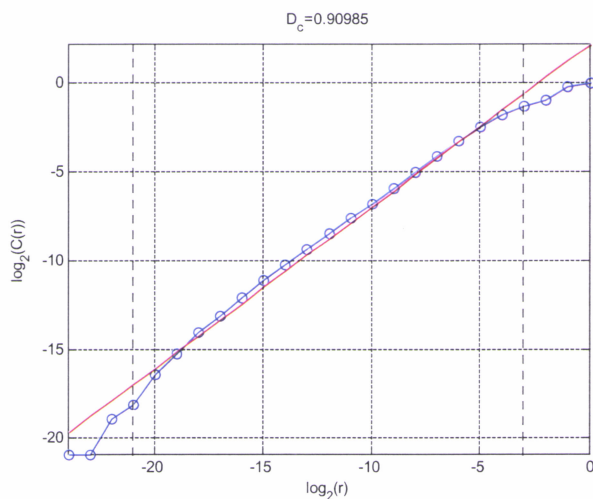


Figura 3.14: Dimensão de correlação para $a = 3.6$.

cos.

Vamos observar o gráfico onde temos o $\log C(r)$ em função de $\log r$ para a função logística para o parâmetro $a = 3.5$, valor no qual o seu comportamento é de período 4.

Como verificamos na figura o gráfico mostra a série em ramos. Se r é menor que a distância mínima entre 4 pontos periódicos, logo $\log C(r) = \log(1/4) = -0.602$ desde que 1/4 dos dados estejam acima um do outro. À medida que r aumenta, será grande o suficiente para que 2 dos 4 pontos periódicos estejam a uma distância menor que r um do outro (para os dados da função logística, os 4 pontos periódicos não estão igualmente espaçados). Para tal valor de r , temos $\log C(r) = \log(1/2)$. Se aumentarmos ainda mais o r o suficiente para incluir todos os quatro valores. Teremos então $\log C(r) = 0$.

Funções no plano Consideremos a função de Hénon

$$x_{n+1} = y_n + 1 - ax^2,$$

$$y_{n+1} = bx_n,$$

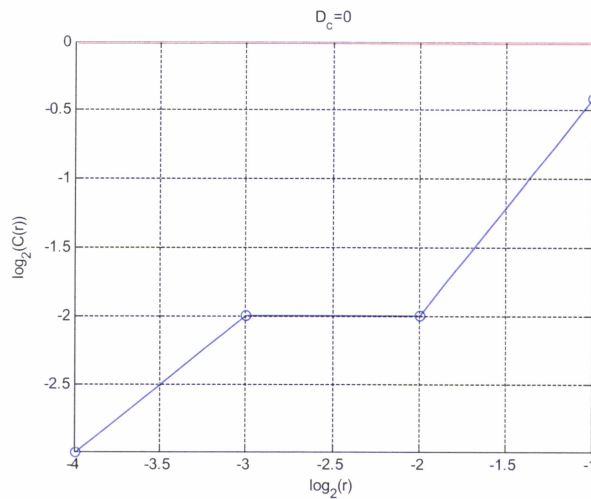


Figura 3.15: Dimensão de Correlação para $a = 3.5$.

com $a = 1.4$ e $b = 0.3$

Muitas vezes não temos acesso a uma série temporal $\{X_n\}$ de vetores F -dimensionais. Em lugar disso temos apenas um ou quanto muito alguns componentes de X_n . Isto é particularmente relevante para experiências reais onde o número de graus de liberdade é frequentemente muito alto se não infinito. Tais sistemas contudo podem ter atratores de baixa dimensão. É então muito desejável possuímos um método seguro que permita uma caracterização desses atratores para uma série temporal de variável única. $\{x_i, i = 1, \dots, N; x_i \in R\}$.

A ideia essencial consiste em construir vetores d -dimensionais

$$\xi_i = (x_i, x_{i+1}, \dots, x_{i+d-1})$$

e usar o espaço dos ξ em vez do espaço dos X . O integral de correlação pode ser, por exemplo

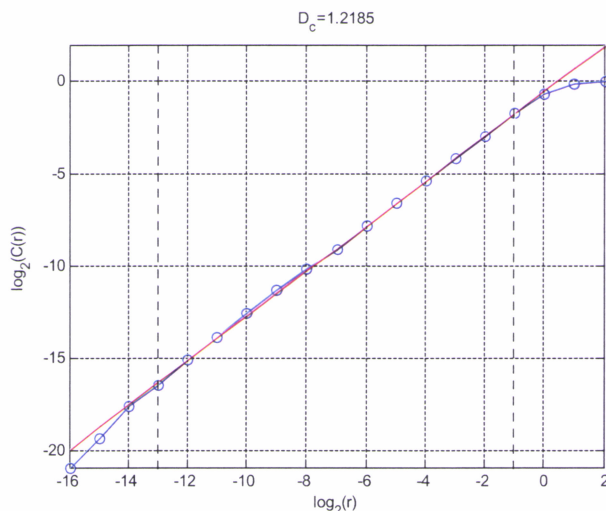


Figura 3.16: Dimensão de Correlação para $a = \frac{1}{4}$ e $b = 0.3$ na função de Hénon.

$$C(l) = \lim_{N \rightarrow \infty} \frac{1}{N(N-1)} \sum_{i,j=1, i \neq j}^N \theta(l - |\xi_i - \xi_j|).$$

Mais geralmente, pode-se usar

$$\xi_i = (x(t_i), x(t_i + \tau), \dots, x(t_i + (d-1)\tau)),$$

com τ fixo nalgum intervalo. A magnitude de τ não deve ser muito pequena visto que doutro modo $x_i \approx x_{i+\tau} \approx x_{i+2\tau} \approx \dots$ e o atrator no espaço dos ξ será esticado na diagonal e assim dificulta a separação. Por outro lado, τ não deve ser muito grande visto que valores distantes nas séries temporais não são fortemente correlacionados.

Um compromisso análogo deve ser tido em conta na escolha da dimensão d . Takens mostrou que se um espaço de estados de um sistema tem dimensão d , então a dimensão de mergulho deve ser pelo menos $2d + 1$ para capturarmos completamente as dinâmicas do sistema.

Por outro lado, não se pode tomar d grande sem provocar erros experimentais e falhas estatísticas.

Vamos agora aplicar estas considerações à função logística com $a = 4$, e à função de Hénon.

Se calculamos para a função logística com $a = 4$ o integral de correlação original e o integral modificado, obtido mergulhando num espaço 2 e 3-dimensional. Observamos na verdade o esperado decréscimo de erros sistemáticos quando aumentamos d , acompanhado por um aumento do erro estatístico.

Para a função de Hénon usamos como séries temporais as séries $\{x_n, x_{n+2}, x_{n+4}, \dots\}$.

Dimensão de Lyapunov Segundo [22], há uma relação estreita entre a **dimensão do atractor** e o **espectro de Lyapunov**. **Kaplan e Yorke** mostraram que a *dimensão de Lyapunov*, D_L , de um atractor pode ser calculada usando os primeiros k expoentes de Lyapunov, com k seleccionado tal que a soma de $(k + 1)$ destes primeiros expoentes seja negativa. D_L é definida como se segue:

Considerando uma órbita com expoentes de Lyapunov $(\lambda_1, \lambda_2, \dots, \lambda_m)$ e com $\lambda_1 > \dots > \lambda_m$ e seja k o maior inteiro tal que $\sum_{m=1}^k \lambda_m \geq 0$.

$$D_L = k + \frac{\sum_{m=1}^k \lambda_m}{|\lambda_{k+1}|} \quad (3.5)$$

Por exemplo, o atractor caótico de Lorenz tem três expoentes de Lyapunov sendo eles, 1.51, 0.0 e -22.5 . Dado que a sua soma se torna negativa com o terceiro expoente, k é um conjunto de 2 e estes valores vão ser substituídos em 3.5 e obtemos $D_L = 2.07$.

Segundo [8], Brown et al. sugeriram que normalmente apenas os primeiros $k + 1$

expoentes de Lyapunov devem ser considerados em qualquer avaliação quantitativa do caos.

E segundo [34], Mori propôs uma fórmula alternativa para estimar a **dimensão fractal** de séries temporais usando o espectro de Lyapunov.

$$D_M = m^0 + m^+ \left(1 + \frac{\lambda^+}{\lambda^-} \right) \quad (3.6)$$

Esta equação é calculada usando m^0 e m^+ , isto é, quantidade de números (aproximadamente) zero e de expoentes de Lyapunov positivos, respectivamente, como se vê em [14].

Como está referido em [22], segundo dados de Gallez e Babloyantz, se os expoentes de Lyapunov positivos são 8.75, 5.31 e 2.15, o (aproximado) zero é -1.63 e os expoentes negativos são -4.52 , -9.65 , -15.17 , e -36.47 , $m^0 = 1$, $m^+ = 3$, $\lambda^+ = 5.40$, e $\lambda^- = 16.45$ (vale a pena referir que o mais negativo pode ser falso enquanto $D_L < 7$, se este valor for removido, $D_M = 5.66$, um valor aproximado de 6.03).

Substituindo estes valores em 3.6 temos $D_M = 4.98$, um valor um pouco menor que o estimado 6.03.

Nos dados experimentais devido ao ruído não é possível calcular os expoentes de Lyapunov usando a definição matemática, referido por Brown et al. em [8]. São necessários métodos mais refinados que façam uso de pontos locais distribuídos na vizinhança de cada ponto atrator. Dado que os expoentes negativos correspondem a regiões estáveis do atrator, o seu cálculo numérico exacto pode ser difícil.

Se a dimensão de mergulho é muito grande, estes expoentes negativos podem ser falsos e instáveis numericamente. Brown et al. em [8] sugeriram que expoentes de Lyapunov

falsos podem ser detectados adicionando uma pequena quantidade de ruído Gaussiano com média zero à série temporal. Expoentes que mudam marcadamente seguindo esta manipulação devem ser ignoradas.

No entanto, para a reconstrução do sistema, é necessário determinar *os valores óptimos da dimensão de imersão m e do atraso temporal ξ* .

Dimensão de mergulho e atraso temporal para reconstruir o atractor Precisamos de encontrar a dimensão de mergulho mínima que nos permita ter as mesmas propriedades geométricas que o atractor original num sistema.

Um dos métodos mais utilizados para determinar m é o *método dos falsos vizinhos mais próximos*, introduzido por Kennel, Brown e Abarbanel.

O método dos falsos vizinhos mais próximos diz respeito à existência de falsos cruzamentos da órbita consigo própria. Estes falsos vizinhos surgem quando o atractor é projectado numa dimensão de imersão menor do que a real. Nesta situação, alguns pontos do atractor estão próximos devido apenas ao falso cruzamento das órbitas. Quando aumentamos a dimensão de imersão, o falso cruzamento desaparece e os mesmos dois pontos têm posições distantes no atractor. Assim, este método baseia-se na procura de uma dimensão mínima para a qual as órbitas não se intersectem. Consiste em encontrar o vizinho mais próximo de cada ponto numa dada dimensão e verificar se estes pontos são ainda vizinhos próximos numa dimensão mais alta. A percentagem de falsos vizinhos mais próximos tende para zero quando se encontra a dimensão óptima.

A ideia básica é determinar se um vizinho é verdadeiro ou falso através da projecção do sistema em uma determinada dimensão. Num espaço muito pequeno ($d < m$) um ponto

do sinal pode ser visto como um vizinho, mesmo não sendo, é chamado de falso vizinho. Quando o espaço está imerso em uma dimensão $d > m$, todos os pontos vizinhos de todas as órbitas são vizinhos verdadeiros.

Segundo Mane-Takens, a condição $m > 2d$, sendo d a dimensão real do atrator, é suficiente para uma boa imersão. No entanto, na prática este valor revela-se por vezes muito alto, o que se traduz em mais dificuldades computacionais, podendo ser obtidas boas imersões com escolhas para m bem menores que $2d$.

Segundo [1], podemos reconstruir o atrator num espaço de mergulho de dimensão $m = 2n + 1$, onde n é a dimensão do espaço de fases do atrator. Alternativamente, se a dimensão de "box counting" do atrator é conhecida então uma estimativa do mergulho mínimo é $m > 2D_B$. De notar que m é um inteiro, no entanto D_B não o é necessariamente. Se apenas sabemos D_C então uma dimensão óptima é simplesmente $m > D_C$.

Segundo [28], uma vez que não se sabe à partida a dimensão de mergulho p , o objectivo de Grassberger-Procaccia na sua análise é encontrar a dimensão D_C .

A solução deste problema é calcular D_C em função dos vários valores de p .



Figura 3.17: Dimensão de correlação D_c em função da dimensão de mergulho.

Um caso idealizado de dimensão de correlação D_C contra a dimensão mergulhada. Os pontos \cdot exibem D_C para séries temporais com atrator de dimensão 2.7, enquanto símbolo

x dá-nos v para ruído-branco aleatório.

Para uma série temporal de um sistema sobre um atrator de dimensão D_C , a dimensão de correlação da série temporal oferece um meio de estimar a dimensão do atrator.

No entanto, para sistemas que não estão num atrator, a interpretação de D_C será mais difícil.

Um caso simples para interpretar D_C é o ruído branco aleatório.

Exemplo 3.1.2 *Considere uma sequência aleatória com ruído branco, tal como no modelo*

$$x_{t+1} = A + \rho x_t.$$

Vemos que quando os dados deste modelo são mergulhados com $p = 2$, eles criam um sólido em forma de gota. Uma vez que este sólido cobre toda a figura com tinta (ou pelo menos tal aconteceria se houvesse muitos mais pontos), este modelo é de dimensão 2 (isto é, $D_C = 2$ quando $p = 2$). Analogamente, um mergulho $p = 3$ irá criar um gota de dimensão 3 que iria cobrir completamente o volume de tridimensional, e logo $D_C = 3$ quando $p = 3$. No caso ideal, para ruído-branco aleatório, $D_C = p$. Onde ideal significa termos uma quantidade infinita de pontos. Obviamente, é impossível haver uma quantidade infinita de dados, mas mesmo para um pequeno conjunto de dados a relação $D_C \approx p$ deverá verificar-se. Uma regra prática é que 10^p pontos são necessários para mostrar que $D_C \approx p$ é verdadeiro para qualquer ponto dado p , logo são necessários 1000 (ou 10^3) pontos para mostrar que $D_C = p$ para $p = 3$.

Em [1], dois métodos se destacam na determinação do atraso temporal (ξ) a função de correlação e o método da informação mútua.

Um dos métodos para determinar o atraso temporal é através da função de correlação, C , que compara dois pontos separados na série temporal com um atraso de ξ , e é definida por:

$$C = \frac{\sum_{i=1}^{N-\xi} (x'_i)(x'_{i+\xi})}{\sum_{i=1}^{N-\xi} (x'_i)^2},$$

onde $x'_i = x_i - \bar{x}_i$. O atraso temporal ξ , para a reconstrução do atrator, é tomado num determinado valor específico de C . Os valores específicos mais usuais são: o primeiro valor de ξ no qual o valor de C é 0.5; o primeiro valor de ξ no qual o valor de C é 0; ou o primeiro ponto de inflexão de C .

De seguida, vamos introduzir o método da informação mútua para determinar o atraso temporal óptimo ξ .

A função de informação mútua do atrator reconstruído é dada por:

$$M = \sum_{i=1}^{N-(m-1)\xi} P(x_i, x_{i+\xi}, x_{i+2\xi}, \dots, x_{i+(m-1)\xi}) \log \left[\frac{P(x_i, x_{i+\xi}, x_{i+2\xi}, \dots, x_{i+(m-1)\xi})}{P(x_i)P(x_{i+\xi})P(x_{i+2\xi}), \dots, P(x_{i+(m-1)\xi})} \right]$$

onde $P(x_i)$ é a probabilidade de x_i , $P(x_i, x_{i+\xi}, x_{i+2\xi}, \dots, x_{i+(m-1)\xi})$ é a probabilidade conjunta de $X_i = (x_i, x_{i+\xi}, x_{i+2\xi}, \dots, x_{i+(m-1)\xi})$. M é a medida estatística que mede a dependência das variáveis umas das outras. Se as variáveis são independentes, então, $P(x_i, x_{i+\xi}, x_{i+2\xi}, \dots, x_{i+(m-1)\xi}) = P(x_i)P(x_{i+\xi})P(x_{i+2\xi}), \dots, P(x_{i+(m-1)\xi})$ e temos então $M = 0$. Neste caso o processo era completamente aleatório, tal como o ruído branco. Caso contrário, uma dependência completa resultaria $M = \infty$. Uma boa escolha do atraso temporal requer que a informação mútua seja mínima. A condição para a escolha do atraso temporal é conhecido como o critério da informação mútua mínima. A reconstru

ção de um atrator de dimensão 2 é no plano $x_i, x_{i+\xi}$, para calcular a informação mútua, o plano é dividido com uma grelha com N_c colunas e N_r linhas. A função densidade de probabilidade para x_i e $x_{i+\xi}$ são o resultado de simplesmente somar os pontos do atrator que estão dentro de cada coluna e linha da grelha e dividir pelo total de pontos do atrator. $P(k)$ e $P(l)$ são as probabilidades da ocorrência do atrator na coluna k na linha l respectivamente. A probabilidade conjunta, $P(k, l)$, de ocorrer o atrator em qualquer caixa particular é calculado por contar o número de pontos discretos na caixa e dividir pelo número total de pontos da trajectória no atrator. A informação mútua encontra-se a partir de:

$$M = \sum_{k=1}^{N_c} \sum_{l=1}^{N_r} P(k, l) \log \left[\frac{P(k, l)}{P(k)P(l)} \right]$$

O valor de ξ que nos dá o primeiro mínimo na função de informação mútua é o atraso temporal na reconstrução do atrator.

O método da informação mútua procura identificar a máxima quantidade de informação que se pode ter de uma medida realizada num determinado tempo i , quando se observa uma outra medida, do mesmo sinal, mas num tempo posterior $i + \xi$. Se ξ é muito pequeno, x_i é semelhante a $x_{i+\xi}$ e quando representados graficamente, os pontos ficam muito próximos de uma linha. Se ξ é muito grande, a correlação entre as coordenadas é muito baixa e não se consegue retirar informação a partir do gráfico.

Com isso, garante-se que o espaço de estado reconstruído seja topologicamente equivalente ao espaço do sistema dinâmico original.

Exemplo 3.1.3 *Estimemos a dimensão de um atrator caótico usando o modelo con-*

hecido como *Função Ikeda*

$$x_{t+1} = 1 + \mu (x_t \cos m_t - y_t \sin m_t),$$

$$y_{t+1} = \mu (x_t \sin m_t - y_t \cos m_t).$$

O primeiro passo é mergulhar a série temporal.

Isto requer a escolha de uma dimensão de mergulho p e um atraso ξ .

Uma maneira de escolher um atraso é tomar o mais pequeno valor de ξ na qual a função de autocorrelação $C \approx 0$. Neste caso, temos que $C(1) \approx 0$, logo vamos usar $\xi = 1$.

Vamos repetir os cálculos para $p = 1, 2, \dots, 10$. Em cada uma destas dimensões de mergulho repetimos os mesmos passos.

1. Calculamos o integral de correlação $C(r)$.

2. Usando a equação $\log C(r) = v \log r + \log A$, usamos $C(r)$ para calcular a dimensão v . Uma maneira de fazer isto é desenhar $\log C(r)$ contra $\log r$, sendo v é o declive deste gráfico $v = \frac{\log C(r)}{\log r}$.

No entanto, este declive geralmente depende do valor de r seleccionado. Para ultrapassar este problema, vamos desenhar o declive v como função de r .

Na figura 3.18 vemos como o declive v muda conforme o r para diferentes dimensões de mergulho.

Para os maiores valores de r encontrados em séries temporais mergulhadas, o declive aproxima-se de zero, independentemente da dimensão mergulhada usada.

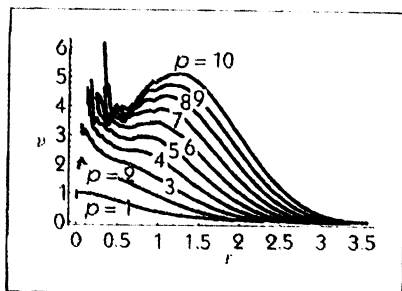


Figura 3.18: Declive v contra r para dimensões mergulhadas desde $p = 1$ a $p = 10$.

Estimar a dimensão se usando um r grande é como olhar para um objecto a uma longa distância. Não importa qual é o objecto, ele parecerá como um ponto singular, um objecto de dimensão 0, tal como as estrelas que aparecem como um ponto de luz.

Para pequenos valores de r , o declive v depende e cresce com a dimensão de mergulho.

É uma característica do ruído e reflete a medida de ruído dos dados deste modelo.

Para $r \approx 0,5$ o declive é aproximadamente o mesmo para diferentes dimensões de mergulho.

A figura 3.19 mostra v (em $r = 0.5$) para $p = 1, 2, \dots, 10$.

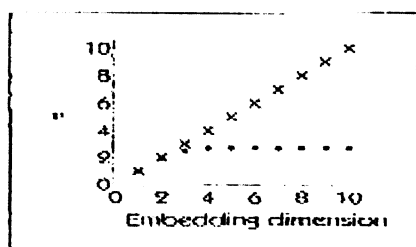


Figura 3.19: Dimensão de correlação versus a dimensão de mergulho para $r = 0.5$.

O padrão é similar ao visto na figura anterior e concluímos que o atractor deste modelo tem a dimensão de aproximadamente 3,9.

O intervalo de valores de r no qual v é praticamente constante é a chamada *região de escala*, como já referimos acima.

A necessidade escolher um intervalo específico de r , é uma das dificuldades, e uma grande fragilidade na estimativa das dimensões. Neste caso, escolhemos um intervalo próximo de $r = 0.5$ porque esse valor dá-nos resultados próximos da forma ideal. Neste caso temos a vantagem da existência de um atrator, logo temos uma boa razão para acreditar que a nossa escolha de r se justifique.

Sem esta informação, interpretar o significado do cálculo da dimensão podia ser alterado e problemático.

Determinismo Em [28], descreveram um método para testar determinismo na série temporal, que é, o que a vai diferenciar da série temporal puramente aleatória.

Dizemos que um sistema é determinístico quando eventos futuros são causados por eventos passados.

A equação $x_{t+1} = f(x_t)$ é determinística enquanto $f(x_t)$ tem apenas um valor para cada possível valor de x_t , dado o valor antecedente de x_t , a função determina o valor futuro x_{t+1} .

Para o modelo da função quadrática, na qual a equação produz caos, se conhecermos x_0 , por iteração podemos calcular todos os valores futuros de x_t usando a equação

$$x_{t+1} = \mu x_t (1 - x_t).$$

Mas, muitas vezes não conhecemos x_0 exactamente porque as nossas medidas têm ruído: $D_t = x_t + w_t$. Se tomarmos D_0 como a nossa estimativa da condição inicial x_0 , e a iterarmos usando a função quadrática, logo a sensibilidade às condições iniciais vai fazer

com que as nossas previsões sejam defeituosas.

Por exemplo, supondo que a condição inicial verdadeiramente é $x_0 = 0.37$ mas as nossas medidas da condição inicial é $D_0 = 0.38$.

Os valores de x_t e y_t das colunas da tabela 3.7 são os valores das equações $x_{t+1} = 4x_t(1 - x_t)$ e $y_{t+1} = 4y_t(1 - y_t)$. As condições iniciais são $x_0 \approx y_0$. A partir da 5ª iterada os valores de x e y distanciam-se um do outro.

x_t	y_t	t
0.370	0.380	0
0.932	0.942	1
0.252	0.217	2
0.754	0.680	3
0.741	0.870	4
0.767	0.451	5
0.715	0.990	6
0.814	0.038	7
0.605	0.147	8
0.956	0.501	9
0.167	0.999	10

(3.7)

Como mostra a tabela 3.7 as nossas iterações baseadas na função logística irão divergir dos valores verdadeiros.

No princípio as iterações serão bastante boas, mas depois de 5 ou mais passos, elas tornam-se completas erradas. Mesmo quando as iterações são boas, elas não são perfeitas porque as medidas têm ruído.

No entanto, é possível decidir através dos dados se um sistema é determinístico.

Usamos os dados para construir um modelo das dinâmicas e logo ver se as iterações saídas deste modelo são exactas.

Se as iterações são perfeitas logo o sistema é completamente determinístico. Se as iterações são boas, mas não perfeitas, logo o sistema tem uma componente determinística.

Se as iterações são más logo o sistema não é determinístico.

Nós podemos construir modelos dinâmicos de dados numéricos de diferentes maneiras.

Supondo que temos medidas até ao tempo T e queremos fazer uma previsão do valor no tempo $T + 1$.

Um dos métodos mais simples é o que se segue:

1. Mergulhar a série temporal para produzir \mathbf{D}_T .
2. Tomar um ponto mergulhado no tempo T .

$\mathbf{D}_T = (D_T, D_{T-h}, \dots, D_{T-(p-1)\xi})$, e olhar para o resto da série temporal mergulhada para encontrar o ponto mais próximo de \mathbf{D}_T . Dizemos que este ponto mais próximo tem índice a . Isto significa que \mathbf{D}_a é mais próximo de \mathbf{D}_T que qualquer outro \mathbf{D}_T .

3. A definição de determinismo é que os eventos futuros são causados por eventos passados. \mathbf{D}_T descreve os eventos anteriores a D_{T+1} . Analogamente \mathbf{D}_a descreve os eventos anteriores a D_{a+1} .

Se \mathbf{D}_T é próximo de \mathbf{D}_a , e se o sistema é determinístico, esperamos que D_{a+1} seja próximo a D_{T+1} .

Logo tomamos como nossa previsão de D_{T+1} o valor da medição D_{a+1} . Chamaremos a isto previsão P_{T+1} .

Os modelos de previsão têm conjuntos consistentes de equações explícitas. Este mo-

delo, o qual é usado para previsão, consiste num conjunto de dados (*as medidas da série temporal*) e um conjunto de instruções dadas (*"encontre o ponto mais próximo de \mathbf{D}_a "*). O conjunto de instruções é chamado um *algoritmo* e o modelo consiste implicitamente num conjunto de dados e o algoritmo.

Tais modelos de dados implícitos eram incomuns antes do advento dos computadores mas agora eles têm o mesmo espaço e são de uma importância crescente.

Há muitas variações sobre este modelo simples das dinâmicas. Uma elaboração é tomar não apenas um tempo a onde \mathbf{D}_a é próximo de \mathbf{D}_T , mas tomar k tempos diferentes a_1, a_2, \dots, a_k onde $\mathbf{D}_{a_1}, \mathbf{D}_{a_2}, \dots, \mathbf{D}_{a_k}$ são todos próximos de \mathbf{D}_T . Logo, a previsão de D_{T+1} é tomar a média de

$$D_{a_1+1}, D_{a_2+1}, \dots, D_{a_k+1} : P_{T+1} = \frac{1}{k} \sum_{i=1}^k D_{a_i+1}.$$

Dado um método para prever P_{T+1} , precisamos de fazer uma medida actual de D_{T+1} em ordem a decidir se a previsão é boa ou má.

A diferença entre P_{T+1} e D_{T+1} é o *erro de previsão*, no qual nos diz sobre a qualidade da previsão. É claro, uma previsão simples talvez seja boa ou má apenas por acaso.

Para dar uma indicação com mais significado do determinismo dos dados, tomamos a média de muitos erros de previsão.

Supomos que fazemos $2T$ medições de uma série temporal. Tomamos a primeira metade da série temporal para construir um modelo de dados implícitos das dinâmicas. Em seguida, usamos o modelo para prever os valores da segunda metade da série temporal.

Há duas maneiras de fazer isto. Uma é usar o modelo para prever o valor no tempo $T + 1$, e depois construir um novo ponto mergulhado usando este valor previsto P_{T+1} :

$$\mathbf{D}_{T+1} = (P_{T+1}, D_{T+1-h}, \dots, D_{T+1-(p-1)\xi}).$$

Continuamos encontrando os pontos próximos de \mathbf{D}_{T+1} e fazendo a previsão do valor no tempo $T + 2$, a qual denotamos por P_{T+2} .

Este processo pode ser iterado - usamos previsões passadas para fazer previsões futuras.

Este método pode de facto ser usado para extrapolar a série temporal além dos seus valores conhecidos à priori.

No entanto, tal extrapolação é útil para fazer previsões num futuro longínquo. Para efeitos de avaliação do determinismo nos dados, é melhor usar os dados medidos directamente.

Nesta segunda maneira de fazer previsões, em ordem a prever o valor no tempo $T + 2$, se tomarmos o ponto mergulhado, $\mathbf{D}_{T+1} = (D_{T+1}, D_{T+1-h}, \dots, D_{T+1-(p-1)\xi})$, note que aqui é usada a medida no tempo $T + 1$, e não a previsão P_{T+1} . Não estamos a usar a previsão passada para fazer previsões futuras.

Uma vez feitas previsões para a segunda metade da série temporal podemos calcular a média do erro de previsão, ε . Com $\varepsilon = \frac{1}{T} \sum_{k=1}^T (D_{T+k} - P_{T+k})^2$. Se ε é muito grande significa que as previsões são más e o sistema não é determinístico. Inversamente ε muito pequeno sugere que o sistema é determinístico.

Supondo que queremos ser mais práticos e que em vez de calcular uma nova previsão para cada tempo $T + K$, fazemos a mesma previsão para todos os tempos. Presumivelmente este é um mau método de previsão, uma vez que ignora completamente as dinâmicas dos dados.



Queremos escolher o valor $P_{prático}$ que minimiza a média do erro de previsão

$$\varepsilon_{prático} = \frac{1}{T} \sum_{k=1}^T (D_{T+k} - P_{prático})^2,$$

a minimização que usamos aqui é idêntica à minimização que desenvolvemos para a encontrar a média simples, M_{est} .

De facto, um bom valor para $P_{prático}$ é a média simples da série temporal.

Dado que fixamos $P_{prático} = M_{est}$, a média do erro de previsão $P_{prático}$ é muito semelhante à variância da série temporal, σ^2 .

Uma maneira conveniente para decidir se ε é grande ou pequeno é compará-lo a $P_{prático}$ ou à variância da série temporal σ^2 . Nós podemos fazer isto tomando a razão $\frac{\varepsilon}{\sigma^2}$.

Se esta razão é aproximada de 1, logo a média do erro de previsão é grande. Se esta razão é próxima de zero, logo a média do erro de previsão é pequena.

Podemos examinar num conjunto de dados de vários modelos à procura de determinismo. Em vez de usar apenas o ponto vizinho mais próximo na série temporal mergulhada para fazermos uma previsão, vamos usar k pontos vizinhos.

Vamos usar uma dimensão de mergulho $p = 1$ logo $\mathbf{D}_t = D_t$. Mas muitas vezes faz sentido escolher $p > 1$.

Para o modelo $x_{t+1} = A + \rho x_t$, podemos colocar num gráfico no eixo das ordenadas a razão entre o erro de previsão e a variância, $\frac{\varepsilon}{\sigma^2}$, e no eixo das abcissas o número de k vizinhos usados para fazer a previsão.

Quando k é pequeno, a razão é maior que um. Isto diz-nos que o modelo faz previsões piores que o método prático de simplesmente prever a média.

Com k cada vez maior a razão aproxima-se da unidade.

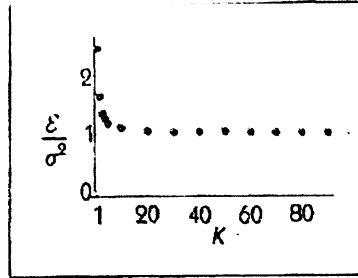


Figura 3.20: Erro de previsão $\frac{\varepsilon}{\sigma^2}$ contra o número de vizinhos k usado na previsão para dados do modelo $x_{t+1} = A + \rho x_t$.

O sistema encontra-se num ponto fixo e toda a variabilidade dos dados deve-se a medidas de ruído aleatório w_t . Não há dinâmicas determinísticas para prever, logo a previsão não é eficiente. Desde que haja 100 pontos na série temporal, quando k é próximo de 100 virtualmente todos os pontos da série temporal estão a ser calculados em conjunto para produzir previsão e a equação $P_{T+1} = \frac{1}{k} \sum_{i=1}^k D_{a_i+1}$ dá-nos a previsão que é basicamente a média, M_{est} . Isto é, portanto, o mesmo que o método de previsão prático de usar a M_{est} e logo $\varepsilon = \sigma^2$. De facto, uma vez que não existem dinâmicas com os dados, nós não podemos fazer melhor que usar M_{est} para prever a série temporal, desde que M_{est} seja a quantidade que minimiza o erro de previsão.

Quando k é pequeno o número de previsão irá variar ao redor da média desde que sejam usados poucos pontos nas médias.

Como é descrito antes, se são usados poucos pontos mais a previsão irá variar ao redor da média. Desde que a média nos dê a melhor previsão possível para esses dados qualquer desvio da média irá dar previsões erradas.

Para o modelo $x_{t+1} = A + \rho x_t$, o erro de previsão é grande, isto é, $\frac{\varepsilon}{\sigma^2} \geq 1$. Temos

uma justificação para concluir que os dados são aleatórios, consistente com o mecanismo conhecido do modelo.

Com o modelo $x_{t+1} = A + \rho x_t + v_t$, em contraste, mostra definitivamente previsibilidade.

A razão $\frac{\varepsilon}{\sigma^2}$ pode ser vista na figura 3.21:

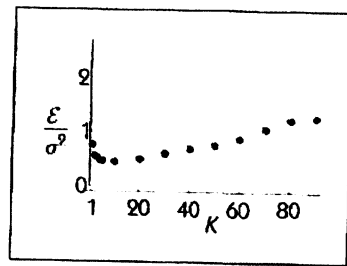


Figura 3.21: Erro de previsão $\frac{\varepsilon}{\sigma^2}$ em função do número de vizinhos k , usado na previsão para dados do modelo $x_{t+1} = A + \rho x_t + v_t$.

Neste modelo a razão $\frac{\varepsilon}{\sigma^2}$ é menor que um para k pequeno e aproxima-se da unidade quando k se aproxima de 100, o número de pontos na série temporal. Para valores muito pequenos de k , as previsões são menos exactas que para valores intermédios.

A média de 5 a 10 pontos próximos produz uma previsão melhor que simplesmente um ponto singular próximo. Esta previsão é substancialmente melhor do que a usada M_{est} .

No entanto concluímos que os dados deste modelo contêm mais determinismo mas não determinismo completo. Isto é consistente com a equação $x_{t+1} = A + \rho x_t + v_t$, na qual tem uma componente determinística nas suas dinâmicas (ρx_t) em soma com ruído aleatório dinâmico (v_t) e medida de ruído (w_t).

O modelo logístico, $x_{t+1} = \mu x_t (1 - x_t)$ é completamente determinístico.

A análise de dados usando erro de previsão confirma isso, o erro de previsão é virtualmente zero para k pequeno e como esperado, a razão $\frac{\varepsilon}{\sigma^2}$ aproxima-se da unidade enquanto

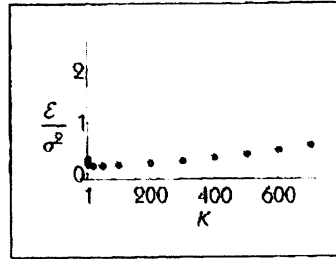


Figura 3.22: Erro de previsão $\frac{\varepsilon}{\sigma^2}$ contra o número de vizinhos k usado na previsão para dados do modelo de duas equações diferenciais associadas $\frac{dx}{dt} = y + v(t)$ e $\frac{dy}{dt} = -ay - bx$ onde $v(t)$ é ruído aleatório.

k se aproxima do número de pontos da série temporal, como se observa na figura 3.23.

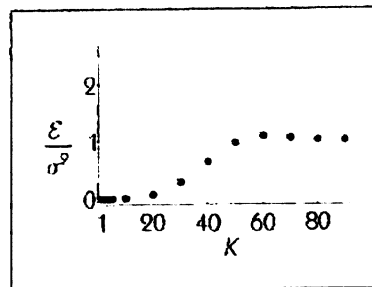


Figura 3.23: Erro de previsão $\frac{\varepsilon}{\sigma^2}$ contra o número de vizinhos k usado na previsão para dados do modelo $x_{t+1} = \mu x_t (1 - x_t)$.

A previsão dos resultados para os modelos $x_{t+1} = A + \rho x_t$, $x_{t+1} = A + \rho x_t + v_t$ e de duas equações diferenciais associadas $\frac{dx}{dt} = y + v(t)$ e $\frac{dy}{dt} = -ay - bx$ onde $v(t)$ é ruído aleatório, poderia ter sido antecipado com a função de correlação.

Segundo [28], a função de correlação C do modelo $x_{t+1} = A + \rho x_t$ mostra que não há nenhuma correlação entre valores medidos D_T e D_{T+1} .

As funções correlação para dados dos modelos $x_{t+1} = A + \rho x_t + v_t$ e de duas equações diferenciais associadas $\frac{dx}{dt} = y + v(t)$ e $\frac{dy}{dt} = -ay - bx$ onde $v(t)$ é ruído aleatório, mostram uma forte co-relação entre D_T e D_{T+1} .

De notar que a função autocorrelação C para o modelo $x_{t+1} = \mu x_t (1 - x_t)$, é a mesma que a do modelo $x_{t+1} = A + \rho x_t$, mas os resultados previstos são completamente diferentes. O método da previsão é sensível à não linearidade do modelo $x_{t+1} = \mu x_t (1 - x_t)$, onde a função de correlação não o é.

O efeito do ruído Atractores de dados experimentais contêm sempre um pequeno elemento aleatório conhecido como o ruído. Este deve-se a flutuações aleatórias no sistema dinâmico (ruído dinâmico), ou a erros aleatórios adicionados ao sinal medido - ruído de medida.

Em [22], Schreiber sugeriu que escolha do atraso e do número da dimensão de mergulho é particularmente importante quando os dados são ruidosos. Usando um atraso relativamente pequeno e uma dimensão de mergulho grande (moderamente), os efeitos de ruído são calculados mais eficientemente. Em qualquer aplicações práticas, os índices quantitativos devem ser calculados para diferentes valores destes parâmetros e assim os seus efeitos podem ser melhor compreendidos.

3.2 Expoentes de Lyapunov

Expoentes de Lyapunov de séries temporais Uma das limitações mais importantes na análise de séries temporais diz respeito à impossibilidade de previsão de um determinado acontecimento. Se um acontecimento é caótico, a sensibilidade às condições iniciais, quantificada pelo maior expoente de Lyapunov, impede a previsão a curto prazo.

Quando as equações que descrevem um sistema dinâmico são conhecidas, os expoentes de Lyapunov podem ser calculados com precisão a partir do algoritmo proposto por Wolf et al. (1985). No entanto, quando se realiza uma experiência, não são medidas todas as variáveis de estado do sistema e o resultado final é uma série temporal que representa a variação de apenas uma variável de estado. Os métodos para a análise dinâmica de séries temporais encontram-se ainda em desenvolvimento. Um método comum consiste em reconstruir o atrator estranho a partir da série temporal e a partir do atrator reconstruído determinar os expoentes de Lyapunov.

Segundo [22], o caos ocorre em sistemas dinâmicos não lineares que exibem dependência sensível às condições iniciais.

Intuitivamente um sistema caótico gera nova informação e incerteza numa escala pequena (localmente) enquanto limita a sua actividade numa grande escala (globalmente).

Consequentemente, a dimensão de correlação fornece-nos um índice quantitativo do comportamento global de um atrator, que necessita ser complementado com outros índices quantitativos que indicam como as órbitas de pontos em atractores divergem e convergem ao longo do tempo. Estes índices medem a incerteza, ou informação, gerada por um sistema dinâmico não linear ao longo do tempo.

Supomos que a trajectória do sistema não linear, mergulhada num espaço de fases de dimensão m , é um pouco perturbada em alguns pontos ao longo do tempo. A relação entre esta pequena perturbação e outra que ocorre num instante posterior depende da taxa de mudança de cada componente da trajectória ao longo do tempo. As relações entre estas pequenas perturbações sobre os pontos do atrator são assumidamente lineares. Logo, é representado por uma tabela, conhecida como a matriz Jacobiana, $J(t)$, contendo as

taxas de mudança para cada coordenada relativamente aos seus valores actuais. O efeito cumulativo que o atractor envolve ao longo do tempo é representado por um produto destas matrizes Jacobianas.

Consideremos o vector, $\mathbf{x}(t) = [x_1(t), x_2(t), \dots, x_i(t), \dots, x_m(t)]^T$, gerado por uma série temporal, $x(t), t = 1, \dots, N$, num espaço de dimensão m , logo o vector no tempo $(t + 1)$ é dada pelo seguinte:

$$\mathbf{x}(t + 1) = \begin{pmatrix} \frac{\partial x_1(t+1)}{\partial x_1(t)} & \frac{\partial x_1(t+1)}{\partial x_2(t)} & \dots & \frac{\partial x_1(t+1)}{\partial x_i(t)} & \dots & \frac{\partial x_1(t+1)}{\partial x_m(t)} \\ \frac{\partial x_2(t+1)}{\partial x_1(t)} & \frac{\partial x_2(t+1)}{\partial x_2(t)} & \dots & \frac{\partial x_2(t+1)}{\partial x_i(t)} & \dots & \frac{\partial x_2(t+1)}{\partial x_m(t)} \\ \dots & \dots & \dots & \dots & \dots & \dots \\ \frac{\partial x_i(t+1)}{\partial x_1(t)} & \frac{\partial x_i(t+1)}{\partial x_2(t)} & \dots & \frac{\partial x_i(t+1)}{\partial x_i(t)} & \dots & \frac{\partial x_i(t+1)}{\partial x_m(t)} \\ \dots & \dots & \dots & \dots & \dots & \dots \\ \frac{\partial x_m(t+1)}{\partial x_1(t)} & \frac{\partial x_m(t+1)}{\partial x_2(t)} & \dots & \frac{\partial x_m(t+1)}{\partial x_i(t)} & \dots & \frac{\partial x_m(t+1)}{\partial x_m(t)} \end{pmatrix} \mathbf{x}(t) \quad (3.8)$$

As entradas da matriz são as derivadas parciais das componentes do vector no tempo $(t + 1)$ relativamente a todas as componentes no tempo t . Esta expressão 3.8 pode ser escrita mais sucintamente como se segue:

$$\mathbf{x}(t + 1) = J(t)\mathbf{x}(t) \quad (3.9)$$

onde $J(t)$ é a matriz Jacobiana no tempo t . A evolução de um vector mergulhado quando o intervalo de tempo varia desde 1 a $t + 1$, é obtido aplicando a expressão 3.9 repetidamente a $x(1)$. Este processo leva-nos ao seguinte:

$$\mathbf{x}(t + 1) = \prod_{i=1}^t J(i)\mathbf{x}(1) \quad (3.10)$$

De acordo com o Teorema Ergódico multiplicativo de Oseledec's, a taxa de dispersão de uma perturbação inicial pequena em cada direcção ortogonal sobre qualquer ponto do atractor é obtida através dos valores próprios do produto das matrizes de 3.10. Os valores da matriz indicam a quantidade de esticamento (valores positivos) e de contração

(valores negativos) que ocorrem em direcções ortogonais no espaço de mergulho quando são aplicadas as transformações lineares representadas pela matriz.

Nesta aplicação, cada valor define um expoente de Lyapunov e o conjunto dos m valores, por ordem decrescente, define o *Espectro* de Lyapunov.

De seguida, podemos referir algumas propriedades qualitativas dos Expoentes de Lyapunov.

Quando um ponto atractor encolhe todos os pontos na sua base de atracção (o conjunto de pontos que terminam no atractor) em direcção a um ponto singular com dimensão 0 em todas as direcções, então os expoentes de Lyapunov são todos negativos.

Para um ciclo-limite, a discrepância inicial entre pontos atractores é mantida inalterável ao longo do tempo como no sistema linear. Se não há esticamento nem encolhimento, os expoentes de Lyapunov são iguais a 0.

Quando os dados são em intervalos de tempo regulares, este processo cíclico pode gerar um expoente de Lyapunov nulo artificialmente, segundo Brown, em [8].

Para um atractor estranho caótico, a incerteza temporal local gera esticamento em algumas direcções mas mantém a estabilidade encolhendo noutras direcções. Logo, o caos é caracterizado por um conjunto de expoentes positivos e negativos.

O maior expoente de Lyapunov é sempre positivo e no caso de hipercaos os primeiros expoentes de Lyapunov são positivos.

Para manter a estabilidade global de um atractor, o total da quantidade de encolhimento deve ser mais que o esticamento do atractor noutras direcções. Logo, um atractor caótico é definido como contendo pelo menos um expoente positivo, com a soma dos expoentes de Lyapunov negativa para garantir que o atractor seja globalmente estável.

Os expoentes positivos expandem o atractor preenchendo o espaço de fases, onde o primeiro expoente negativo contrai o atractor levando a uma região do espaço de fases.

Uma maneira intuitiva de compreender as dinâmicas e o significado geométrico de um atractor, é visualizar por exemplo, um enxame de insectos. Os insectos aparentemente voam individualmente sem rumo, mas o conjunto globalmente é estável e move-se como uma estrutura.

O maior expoente de Lyapunov mede a divergência exponencial de trajectórias inicialmente próximas no atractor na direcção correspondente ao maior valor próprio. Este expoente pode ser calculado através de uma série temporal escolhendo trajectórias com curtos períodos de tempo.

Os precursores destas técnicas, devem-se a Wolf, Swift, Swinney, e Vastano, ver [46], escolhendo as trajectórias até a sua separação exceder um valor máximo pré-definido. Quando isto ocorre, é escolhida outra trajectória, para minimizar o efeito cumulativo do ruído.

Um exemplo interessante de como falsos expoentes de Lyapunov podem surgir quando a série temporal é mergulhada num espaço de mergulho muito grande como se observa na tabela 3.11. Onde temos as estimativas dos Expoentes de Lyapunov para a função de Hénon em função da dimensão de mergulho. Para a função de Hénon é necessário apenas um mergulho de dimensão 2.

m	λ_1	λ_2	λ_3	λ_4	λ_5	λ_6
2	0.445	-1.609				
3	0.441	0.893	-1.654			
4	0.442	-0.307	-0.804	-1.623		
5	0.463	-0.049	-0.390	-0.761	-1.635	
6	0.465	0.144	-0.227	-0.425	-0.872	-1.645

(3.11)

Estimar os expoentes de Lyapunov depende do número da dimensão de mergulho, m , a estimativa de D_L cresce desde o valor 1.28 à medida que m cresce.

É importante seleccionar um valor apropriado de k , o número de passos para o qual a trajectória deve ser monitorada para estimar a divergência de trajectórias inicialmente próximas. No entanto, na teoria devemos de ser capazes de acompanhar a evolução temporal da trajectória para sempre, na prática isto pode ser feito apenas para um período de tempo finito. Erros numéricos, ou os efeitos do ruído geram instabilidade se as trajectórias se cruzam. Brown et al., em [8] sugeriu que k é inversamente proporcional à frequência dominante no espectro da série temporal.

De acordo com Schreiber, em [43] estimar o expoente máximo de Lyapunov usando dados experimentais é difícil. A vantagem de usar programas computacionais quando aplicamos o algoritmo proposto por Rosenstein, Collins, e De Luca, como se vê em [42] é que calcula com eficiência o expoente máximo de Lyapunov para um conjunto de dados pequenos, não contendo mais que 500 observações. Este algoritmo calcula distâncias entre

pontos vizinhos numa série temporal mergulhada que estão separados por, pelo menos, na série temporal, por um período médio, dado pelo inverso do cálculo das componentes do espectro.

Para cada vector de mergulho na série temporal, $x(i)$, é determinado o seu vizinho, $x'(i)$. A evolução temporal da distância entre estes pontos é registada para vários pontos iniciais da série temporal. Quando escolhemos t períodos de tempo, as distâncias entre pares de pontos são registados pelo seguinte:

$$d_i(t) = d_i(0) e^{\lambda_{\max} t} \quad (3.12)$$

onde $d_i(t)$ é a distância entre o i -ésimo par de pontos inicialmente próximos após t passos temporais e $d_i(0)$ é a distância inicial desses pontos. Tomando logaritmos em ambos os membros da 3.12 temos o seguinte:

$$\ln[d_i(t)] = \ln[d_i(0)] + \lambda_{\max} t \quad (3.13)$$

O máximo expoente de Lyapunov, λ_{\max} , é estimado calculando os declives obtidos pela regressão linear de $\ln[d_i(t)]$ com o tempo, t , para um número de vectores iniciais, $\mathbf{x}(i)$, usando a 3.13. Este procedimento fornece-nos uma estimativa razoável de λ_{\max} que não é afectada por ruído. Caos é indicado pela invariância do maior expoente de Lyapunov quando o número de dimensões mergulhadas é alterado.

Bask em [3] aplicou o método de Rosenstein et al.'s(1993) para estimar o maior expoente de Lyapunov para algumas séries temporais alteradas, usando a técnica de Rosenstein et al. para estimar o atraso de mergulho. Esta técnica calcula a distância de sepa-

ração de cada componente do ponto na qual pontos próximos no espaço tendem a divergir ao longo do tempo. É obtido um atraso de mergulho apropriado quando esta medida atinge um valor constante.

Schreiber, em [43] sugeriu que as técnicas simples para calcular os expoentes de Lyapunov, como aqueles que usam a taxa média de variação ao longo do tempo de um pequeno volume contendo pontos atratores (Wolf et al.,1985), não são fiáveis. Isto é verdadeiro para dados experimentais, a menos que haja razões para suspeitar que o processo dinâmico seja determinístico.

Um algoritmo alternativo, deve-se a Kantz, em [26] e explicado em Schreiber, em [43] anula flutuações de ruído para estimar com exactidão a taxa de variação exponencial da separação entre dois pontos inicialmente próximos do atractor. O maior expoente de Lyapunov é estimado pelo declive do gráfico em que num eixo está o tempo e no outro a soma dos logaritmos da divergência média entre pontos inicialmente próximos.

Num exemplo da aplicação desta técnica, Schreiber, em [43] mostrou que é difícil encontrar uma região estável para estimar o maior expoente de Lyapunov para intervalos de dados dos batimentos cardíacos num humano.

Ao contrário da maioria, Bask, em [3] usou um método semelhante para calcular com intervalo de 95% de confiança o maior expoente de Lyapunov, λ_{\max} , de modo que hipóteses nulas que $\lambda_{\max} = 0$ possam ser avaliadas. O método estima os parâmetros e seus erros padrão usando partes diferentes do mesmo conjunto de dados experimentais sem recurso à teoria da estatística tradicional. Se as hipóteses nulas são rejeitadas e o valor λ_{\max} dos dados experimentais é positivo, logo o caos deve estar presente.

Golia e Sandri, em [19] usaram um método semelhante para calcular simples erros no

cálculo dos índices quantitativos tais como a dimensão correlação D_C e o maior expoente de Lyapunov. Eles estimaram uma forma funcional desconhecida para terem a possibilidade nos sistemas dinâmicos não lineares usarem redes neuronais artificiais. As entradas das redes neuronais tem k pontos de falsos vizinhos da série temporal em análise, depois de as séries serem mergulhadas num espaço de dimensão adequado. As saídas das redes são a evolução temporal desses falsos vizinhos T períodos de tempo para o futuro.

A aplicação não linear, estimada pelas Redes Neuronais, é usada para gerar simples pontos da série temporal. Analogamente, geram-se séries temporais diferentes usando pontos de partida seleccionados aleatoriamente para o algoritmo de previsão não linear.

Usando esta técnica, o valor médio para o maior expoente de Lyapunov para a série temporal de Lorenz é 0.6924 (SD=0.0184). É favorável com o valor teórico de 0.6932.

Os expoentes de Lyapunov Globais representam a tendência na qual o atractor diverge ou contrai em direcções ortogonais, determinado pelos valores gerados pela taxa de variação de trajectórias inicialmente próximas. Ao contrário, expoentes de Lyapunov locais, calculado usando a técnica dos falsos vizinhos, medem a mudança temporal para pontos individuais dentro de um atractor. Por esta razão os expoentes de Lyapunov locais devem ser bastante não homogéneos para regiões diferentes do atractor (Bask,1998).

Gençay e Dechert, em [17] propuseram um método para estimar o espectro de Lyapunov Global da série temporal usando redes neuronais para prever a próxima observação da série, usando as L observações prévias, número de neurónios na camada. É usada uma rede neuronal artificial multicamada para prever valores futuros nas séries temporais.

Estas redes neuronais têm m unidades de entrada (m é a dimensionalidade do mergulho no espaço), L unidades e 1 unidade de saída.

A função não linear aplica o vector inicial $z = [z_i]$, $i = 1, \dots, m$ numa saída com um escalar simples, v , dado pelo seguinte.

$$v = \sum_{j=1}^L \beta_j f \left(\sum_{i=1}^m w_{ij} z_i + b_j \right) \quad (3.14)$$

onde β_j , $j = 1, \dots, L$, é o peso de conexão entre a j -ésima unidade e a unidade de saída; w_{ij} , $i = 1, \dots, m$, $j = 1, \dots, L$, é o peso da conexão entre a i -ésima unidade de entrada e a j -ésima unidade e b_j , $j = 1, \dots, L$ são as bias (unidade especial usada para aumentar os graus de liberdade, permitindo uma melhor adaptação, por parte da rede neural, ao conhecimento a ela fornecido) para a j -ésima unidade. A função $f(x)$ é a função não linear sigmóide dada pelo seguinte:

$$f(x) = \frac{p}{1 + e^{-(qx+r)}} \quad (3.15)$$

onde p , q e r são constantes. A vantagem de usar redes neuronais é que podemos aproximar qualquer função contínua não linear com uma precisão arbitrária, desde que se use unidades suficientes (Hornik, Stinchcombe, e White, 1990).

Uma vez que a função foi estimada, as derivadas podem ser calculadas e o espectro de Lyapunov pode ser estimado usando a sua definição.

Gençay e Dechert, em [17] mostraram que redes neuronais fornecem-nos estimativas acertadas dos expoentes de Lyapunov para a função logística (estimado em 0.673 e na teoria é de 0.674) e a de Hénon (estimados em 0.408 e -1.620 e na teoria são de 0.405 e -1.625), usando simples observações de 100 a 200, respectivamente. Gençay e Dechert (1992) forneceram igualmente impressionantes trabalhos com redes neuronais no modelo do atrator de Lorenz, desta vez com 2000 observações. As estimadas da rede neuronal não são afectadas até 10% da medida do ruído e 0.7% sistema de ruído.

Quando a série temporal de Lorenz é mergulhada em quatro em vez de três dimensões, é falso o maior expoente de Lyapunov negativo. Todavia, como Gençay e Dechert (1992) notaram, que o expoente real de Lyapunov deve ser relativamente invariante enquanto cresce o número de dimensões. Por esta razão, deve ser evitado usar um número elevado de dimensões quando são analisados dados experimentais.

Gençay e Dechert, em [17] sugeriram que uma estimativa acertada para o cálculo dos expoentes de Lyapunov requiere um mergulho de dimensionalidade igual à dimensionalidade do atrator, sendo um requisito impossível para séries temporais experimentais. Se o atrator é mergulhado num grande número de dimensões, logo qualquer expoente que excede pode ser falso. Em alguns casos o maior destes expoentes falsos excede o maior expoente de Lyapunov verdadeiro. Dado que a magnitude do maior expoente de Lyapunov quantifica uma quantidade de caos na série temporal, deve existir caos.

Expoentes falsos são estimados quando as dimensões do mergulho no espaço é maior que a apropriada para representar o atrator. Este problema pode ser resolvido assumindo que o número de expoentes verdadeiros é igual ao menor inteiro superior à dimensão fractal do atrator, como D_2 . Expoentes falsos podem ser detectados calculando o espectro de Lyapunov usando a versão do tempo invertido da série temporal. Ao contrário dos expoentes verdadeiros, os expoentes falsos não mudam o sinal com o tempo invertido. Para ultrapassarmos algumas destas dificuldades, os expoentes de Lyapunov devem ser calculados usando o espaço tangente do atrator. Este espaço representa a taxa de variação local de pontos atratores. Usando este método, os expoentes de Lyapunov são calculados

como valores próprios da matriz de dimensão $m \times m$:

$$D = \begin{bmatrix} 0 & 1 & 0 & \dots & 0 & 0 \\ 0 & 0 & 1 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & 0 & 1 \\ v_1 & v_2 & v_3 & \dots & v_{m-1} & v_m \end{bmatrix} \quad (3.16)$$

onde m é o número de dimensões de mergulho. v é a função não linear respeitantes ao $(t + m)$ -iésimo escalar da observação, $y(t + m)$, às m observações precedentes.

$$y(t + m) = v[y(t), y(t + 1), \dots, y(t + m - 1)] \quad (3.17)$$

e

$$v_i = \frac{\partial v}{\partial y(t + i - 1)}, i = 1, \dots, m \quad (3.18)$$

A função v é estimada usando redes neuronais e v_i é o declive da tangente desta função na $(t + i - 1)$ -ésima observação.

Gençay e Dechert, em [17] ilustraram o seu método usando séries temporais geradas pela recursão de Hénon. Para ilustrar os seus propósitos eles transformaram as saídas das séries temporais, $x(t)$, por uma função, $y(t)$, a qual calcula a autoregressão linear da observação corrente e o múltiplo da observação anterior, usando o seguinte:

$$y(t) = x(t) + \alpha x(t - 1) \quad (3.19)$$

O coeficiente de autoregressão, α , num intervalo de 0 a 4.0. Em aplicações psicológicas, este modelo talvez represente um observador que mantém as observações prévias e correntes num trabalho de memória enquanto desenvolve uma tarefa. Quando α é 0.75 ou superior, é gerado o segundo expoente positivo de Lyapunov falso, sugerindo erradamente que o atrator de Hénon é hipercaótico.

Gençay, em [15] mostrou como redes neuronais "multilayer" podem representar as dinâmicas de um sistema não linear. Nesta aplicação, a rede neural aproxima a função não linear e suas primeiras derivadas com mais exactidão que um modelo polinomial não linear baseado na expansão de Taylor (matematicamente semelhante ao sistema de identificação não linear). São necessárias estimativas exactas das primeiras derivadas para estimar os expoentes de Lyapunov. As redes neuronais fornecem-nos uma estimativa consistente da função não linear, porque a estimativa de erro decresce com o crescente número de observações.

Gençay, em [16] também propôs a inicialização, técnica que calcula intervalos de confiança para estimar o espectro de Lyapunov obtido por métodos de previsão de redes neuronais. Tal informação da estimativa exacta é importante, especialmente quando a estimativa é próxima de zero. Usando dados simples para calcular estimativas diferentes do mesmo parâmetro, tais como o maior expoente de Lyapunov, logo 95% dos intervalos de confiança podem ser calculados. Se a média é positiva e o intervalo de confiança não contém zero logo a população do expoente de Lyapunov máximo é positivo, um indicador possível de caos. Gençay demonstrou o sucesso desta técnica usando as séries temporais caóticas conhecidas.

Bask e Gençay, em [3] procederam de forma semelhante para estimar com intervalo de 95% de confiança o maior expoente de Lyapunov e demonstraram a validade do procedimento usando o atractor de Hénon. Bask (1998) usou um procedimento similar para testar hipóteses nulas, que um dos expoentes de Lyapunov é igual a zero. Como os sistemas contínuos, gerados por equações diferenciais, têm pelo menos um expoente de Lyapunov zero, onde as aplicações discretas não têm. Esse teste pode determinar se a dinâmica

subjacente experimentalmente gerada por um atrator surge de um sistema contínuo ou discreto.

Entropia Aproximada Uma medida geral útil de dinâmicas não lineares é entropia aproximada (ApEn), a qual mede quanto a série temporal é imprevisível.

ApEn tem um baixo valor quando as observações tendem a ser seguidas por outras séries com observações similares.

É obtido um valor alto quando a série é mais variável, ou desordenada. Um algoritmo para calcular ApEn foi fornecido por Kaplan et al. (1991). Os procedimentos matemáticos envolvem detectar padrões similares em séries temporais e depois calcular o logaritmo da probabilidade de seguindo observações elas diferem do padrão observado anteriormente. Nesta análise, o comprimento do padrão e o parâmetro do critério semelhante pode ser manipulado (Ho et al. 1997).

Os pormenores técnicos do índice de ApEn estiveram presentes em Pincus (1995). Usando um comprimento máximo, m , e a tolerância, r , a sequência de $N - m + 1$ vectores é gerada da série temporal, cada vector contém m observações sucessivas. N é o número total de observações na série temporal. Para cada vector, é calculada a proporção de vectores que são inferiores a r , usando a distância métrica da diferença máxima entre cada componente do vector. Se $\Phi(m, r, N)$ é o valor médio do logaritmo desta proporção sobre todos os vectores, logo $ApEn$ é definida pelo seguinte:

$$ApEn(m, r) = \lim_{N \rightarrow \infty} [\Phi(m, r, N) - \Phi(m + 1, r, N)] \quad (3.20)$$

Em aplicações práticas, m é 2 ou 3 e r está entre 0.1 e 0.25 do desvio-padrão da série

temporal.

Nos exemplos seguintes m é igual a 2 e r é igual a qualquer um dos destes números 2 ou 0.15 do desvio-padrão, qualquer um é grande. Diferente de muitos índices quantitativos de dinâmicas não lineares, $ApEn$ pode ser calculado com apenas 100 pontos.

$ApEn$ é estimado usando um programa APEN já usado por Pincus, Gladstone e Ehrenkrantz (1991). Os valores para os dados EEG de um gato são $ApEn(2, 2.0) = 0.93$, valores para os dados ao dactilografar IKT são $ApEn(2, 9.17) = 1.82$ e valores para o atrator de Hénon são $ApEn(2, 2.0) = 0.25$.

Claramente o que tem a estimativa de $ApEn$ mais baixa tem dados mais determinísticos e previsíveis.

Outra medida útil de não linearidade para séries temporais é fornecida por análise da flutuação de tendência (DFA), revela-nos o grau de correlações num grande intervalo da série temporal. Um procedimento para calcular este índice é descrito em Peng, Havlin, Stanley e Goldberger (1995). DFA é calculado primeiro por subtrair a média para cada valor, $X(i)$, $i = 1, N$, e depois calcular a soma destes valores para obter $Y(k)$ como se segue:

$$Y(k) = \sum_{i=1}^k [X(i) - \bar{X}] \quad (3.21)$$

A série temporal é dividida em subséries de comprimento igual cada contendo n observações sucessivas. Logo a melhor linha de tendência é calculada para soma destas subséries usando regressão linear. A componente de tendência é subtraída integrando a série temporal em cada segmento e desvio-padrão destes valores residuais é calculado para a série temporal completa usando Ho. et al. (1997)

$$F(n) = \sqrt{\frac{1}{N} \sum_{k=1}^N [Y(k) - Y_n(k)]^2} \quad (3.22)$$

onde $Y(k)$ representa a tendência da série temporal para a k -ésima observação.

Este cálculo é repetido para valores diferentes de n e a relação entre $F(n)$ e n é avaliada graficamente. Quando estes valores são aplicados na função logarítmica representarmos num gráfico este deve ser linear.

Aplicando o teste DFA aos dados da série temporal dos batimentos cardíacos, Ho et al. (1997) este forneceu dois segmentos lineares para a relação entre $\log F(n)$ e $\log(n)$.

O declive do gráfico para um baixo n indica a magnitude de correlações num curto tempo na série temporal, onde o declive para para um n grande indica a magnitude de correlações num período longo.

Foram feitas análises DFA usando o programa DFA para as séries temporais ao IKT, o EEG do gato e a de Hénon.

Os gráficos que relacionam $\log F(n)$ com $\log n$ são lineares com declives 0.611, 0.463, e 0.337 para estas três séries temporais, respectivamente. Uma boa forma linear é vista na figura, a qual mostra o resultado do cálculo de DFA para os dados IKT. Declives perto de 0.5 sugerem que os dados têm bastante ruído.

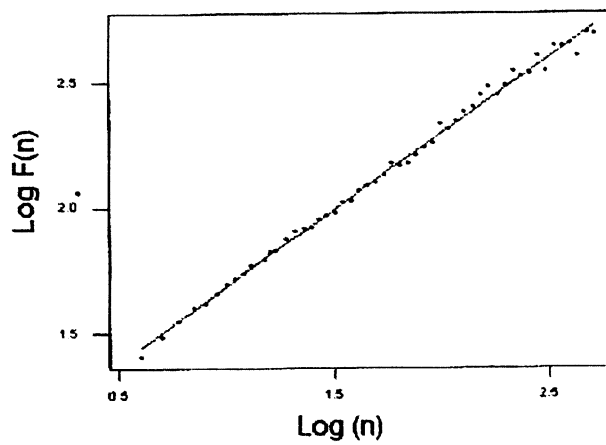


Figura 3.24: DFA para dados ao IKT. O declive α é igual a 0.611.

Capítulo 4

Aplicações à Psicologia

Uma aplicação pioneira das dinâmicas não lineares em Psicologia foi a representação de Gregson de vários fenómenos psico-físicos em termos de recursão cúbica complexa de variáveis de estado, Y_j .

A função Γ recursiva foi desenvolvida, no entanto, para fornecer um modelo dos processos psico-físicos e é baseada na função logística:

$$\Gamma : Y_{j+1} = -a(Y_j - 1)(Y_j - ie)(Y_j + ie).$$

A força da resposta, Y , no tempo $j + 1$ é uma função da resposta do tempo anterior j , a é um parâmetro de controlo representando o comprimento do sinal físico, e é um parâmetro de controlo interno e $i = \sqrt{-1}$ a unidade imaginária. Note que quando o comprimento do sinal se torna suficientemente fraco, a resposta é lançada num regime caótico.

A componente real de Y_j representa a resposta emitida através de um sistema não linear e é calculado com técnicas psico-físicas. A componente imaginária de Y_j não é observada experimentalmente mas é necessária para manter uma estabilidade interna do sistema psico-físico.

Em [22], Gregson referiu que este modelo tem sido usado com sucesso para representar

fenómenos como curvas psico-métricas, interações sociais, transformações na percepção de misturas de padrões, incluindo algumas ilusões geométricas.

Esta equação pode ser expandida a duas respostas e pode tornar-se útil para modelar respostas com modalidade transversal, tal como interpretar a matiz de uma luz colorida através apenas da percepção do brilho ou estudar a ilusão peso-comprimento.

A saída a duas dimensões é realizada pelo seguinte:

$$X_{j+1} = a(e^2 - e^2 X_j + X_j^2 - Y_j^2 - 3X_j Y_j^2 - X_j^3)$$

$$Y_{j+1} = aY_j(-e^2 + 2X_j - 3X_j^2 + Y_j^2).$$

Todos os processos psíco-físicos, segundo Gregson, são o resultado de pelo menos duas funções, uma envolvendo a relação entre estímulo e resposta, e uma referente ao fornecimento dos estímulos ao longo do tempo. Virtualmente todos os estímulos em experiências convencionais utilizam um intervalo fixo. Tal regularidade não pode ser assumida no mundo real.

A teoria de sistemas dinâmicos não lineares gerou inovações na teoria da aprendizagem. Neuringer and Voss (1993) continuaram os trabalhos de Estes relativamente à teoria da aprendizagem estatística, para avaliar se os humanos podem aprender a reconhecer padrões caóticos e fazer previsões de um elemento seguinte numa série de estímulos. Fizeram-se três experiências. A primeira com quatro objectos que estão eventualmente sobre um gráfico de linhas de uma função caótica, baseada na seguinte equação logística:

$$y_{j+1} = 4y_j(1 - y_j)$$

A tarefa é colocar o cursor na posição do próximo ponto antes do programa. Após sucessivas provas, a diferença entre o ponto posicionado e os valores fornecidos pelo programa decresceu para um mínimo assintótico numa curva de competências de aquisição da aprendizagem.

Em [22], Babloyantz e Destexhe examinaram as propriedades dinâmicas do registo de um EEG (electroencefalograma) simples durante uma crise epiléptica. O gráfico do atractor em três dimensões é estruturado com $D_C = 2.05 \pm 0.09$. O expoente de Lyapunov máximo de 2.9 ± 0.6 indicando que a série temporal do EEG é caótica. Os registos durante o sono no estágio 4 revelam um atractor com $D_C = 4.05 \pm 0.05$, e durante o estágio 2 $D_C = 5.03 \pm 0.07$. Foram obtidos valores semelhantes para D_C para estes dois estágios de sono por Babloyantz, Salazar, e Nicolis. O maior expoente de Lyapunov positivo para cada estágio de sono indica que a série temporal seja caótica. Nestes dois estudos, o valor de D_C não pode ser calculado em registos de EEG num estado desperto, presumivelmente devido à coerência reduzida na actividade cerebral e a possíveis efeitos das medidas de ruído. Um ano depois, obtiveram $D_C = 6.1 \pm 0.5$ para estágios alfa relaxado e um valor superior $D_C = 9.7 \pm 0.7$ para um estágio beta activo.

Gallez e Babloyantz sugeriram que D_C para um registo simples de EEG de pacientes Creutzfeldt-Jacob¹ é cerca de 3.8. O espectro de Lyapunov é caracterizado por dois expoentes positivos e um expoente suficientemente grande negativo, logo a soma dos expoentes de Lyapunov é negativa. Por isso, os registos de EEG da doença de Creutzfeldt-Jacob, bem como, para o sono profundo e registos rítmicos alfa, é provavelmente caótica.

¹A doença de Creutzfeldt-Jacob (DCJ) é uma desordem cerebral caracterizada por perda de memória, tremores, desordem na marcha, postura rígida e ataques epilépticos e paralisia facial devida a uma rápida perda de células cerebrais causada por uma proteína transmissível chamada príon.

Os valores de D_C nos registos de uma série temporal de um EEG durante o sono profundo e no estado alfa relaxado são 4.4 e 6.1, respectivamente. Durante o relaxamento, três expoentes de Lyapunov positivos indicam a presença de hipercaos. Babloyantz confirmou este resultado mostrando que o gráfico de recorrência do EEG de pacientes com Creutzfeldt-Jacob é mais estruturado comparado com participantes saudáveis em sono profundo.

Estes autores foram também particularmente cuidadosos no cálculo dos expoentes de Lyapunov. Verificaram a invariância das estimativas quando o número de pontos, a evolução temporal de duas trajectórias inicialmente próximas, e a dimensão de mergulho, assim como a frequência do EEG foram variados. Repararam que um valor elevado da dimensão de mergulho gera um falso expoente de Lyapunov negativo. Quando os expoentes de Lyapunov para EEG de ritmos alfa são comparados individualmente e com os outros participantes, há diferenças no maior expoente de Lyapunov entre participantes, mas os expoentes de Lyapunov negativos são relativamente constantes. Para qualquer participante, o maior expoente de Lyapunov é surpreendentemente estável.

A entropia Kolmogorov-Sinai é mais baixa para Creutzfeldt-Jacob e registos de sono profundo, aproximadamente 7, do que para registos de EEG no estado alfa, cerca de 20 em média. No entanto, os valores mais baixos dependem da taxa de amostragem. Quando a dimensionalidade fractal dos registos do EEG é calculada usando a fórmula de Kaplan-Yorke, os valores são semelhantes, mas ligeiramente superiores que D_C . Gallez e Babloyantz concluíram que o EEG é mais provável que seja caótico. Provaram ainda que o atractor para a série temporal de EEG para Creutzfeldt-Jacob contém duas dinâmicas distintas, um estado caótico e um hipercaótico. Suas técnicas são úteis quando dois ou

mais estados mentais governam o comportamento.

Gregson et al. calcularam D_C usando registos simples de um EEG de participantes relaxados prevendo uma sequência de luzes. O conteúdo da informação na tarefa de previsão das luzes é controlada por aumentar o número de luzes monitorizadas desde 1 a 3. Os dados do EEG foram mergulhados num espaço de dimensão 23 com um atraso temporal de 8. E assim, calcularam cuidadosamente D_C usando pares aleatórios de pontos no atractor. O resultado de D_C foi de 7.98 com os participantes de olhos fechados e D_C aumenta com o conteúdo da informação da tarefa de previsão das luzes para apenas dois de seis participantes.

A dimensão de correlação, D_C , tem sido aplicada num grande conjunto de situações matemáticas. Nesta secção a D_C , é calculada para alguns conjuntos de dados pequenos e com ruído, no entanto, é necessário uma análise mais aprofundada seguindo procedimentos de redução de ruído.

De seguida, vamos observar uma aplicação da dimensão de correlação em dados temporais adquiridos ao digitar uma informação numa imprensa.

Antes de calcular os índices não lineares, é importante analisar as propriedades lineares dos seus dados e verificar a tendência dos dados para vermos se a série temporal não é estacionária. Para estes dados há uma tendência linear significativa estatisticamente dada pela equação seguinte:

$$f(n) = 170.81 + 0.00097n$$

Como se vê na figura 4.1, quando nos dois eixos, o espectro em função da frequência são aplicados os logaritmos há um declive negativo igual a 0.216 ($R^2 = 0.368, p < 0.001$). A regressão é a linha contínua e a linha a tracejado corresponde ao intervalo de confiança

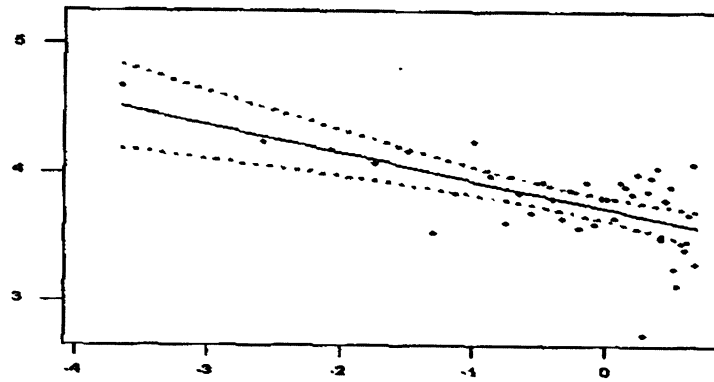


Figura 4.1: Gráfico de dados temporais da tendência ao dactilografar.

de 95%.

A função de autocorrelação que atingiu o ponto zero no valor 4.

A tarefa seguinte é obter valores apropriados para o número mínimo da dimensão de mergulho, estimado pelo cálculo da proporção dos falsos vizinhos enquanto o valor da dimensão mergulhada aumenta. Esta proporção é perto de 0 quando há já uma dimensão suficiente para vermos adequadamente o atractor no espaço de fases.

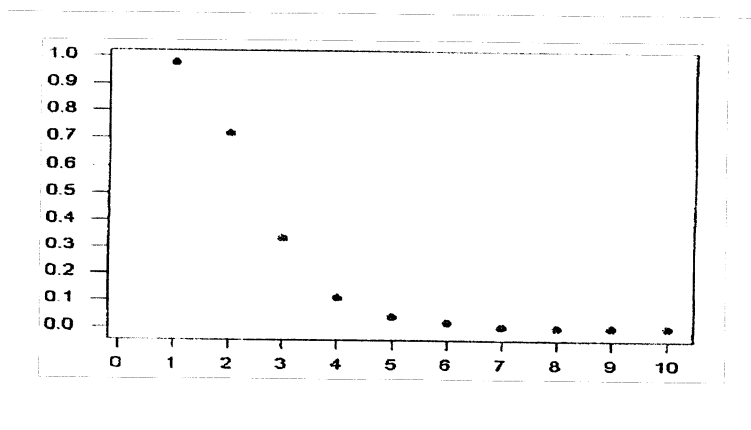


Figura 4.2: Proporção dos falsos vizinhos em função do número das dimensões de mergulho.

Como se vê na figura 4.2 a proporção dos falsos vizinhos aproxima-se de 0 na vizinhança da dimensão de mergulho 5. Logo, D_C , pode ser devidamente interpretado com pelo menos uma dimensão de mergulho 5.

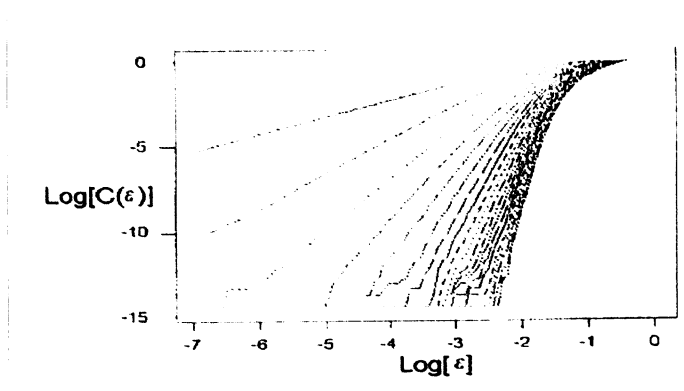


Figura 4.3: Temos o logaritmo do Integral de Correlação, $\log(C_r)$ em função do logaritmo de r , $\log r$ para os dados ao dactilografar.

O gráfico da figura 4.3 mostra que tem segmentos lineares para pequenos m , mas não tem segmentos lineares discerníveis para grandes m . Quando estimamos o declive no intervalo entre -2.3 e -1.7 , este cresce linearmente. Este declive crescente indica uma falta de determinismo não linear no conjunto dos dados, é característica de ruído. Estes dados devem ser revistos após a redução de ruído.

De seguida, vamos analisar os dados do ritmo cardíaco, estes dados sobre os batimentos cardíacos foram registados durante a realização de uma tarefa. Os dados foram registados originalmente por Gibson (1995) e analisados por duas maneiras: Pela filtragem de Kalman e pelas técnicas das dinâmicas não lineares por Heath, Mascord, Gibson e Starmer(1996). Cada uma das 1364 observações do conjunto de dados foi obtida registando intervalos entre dois batimentos cardíacos durante um período de observação de 10 segundos. A tendência

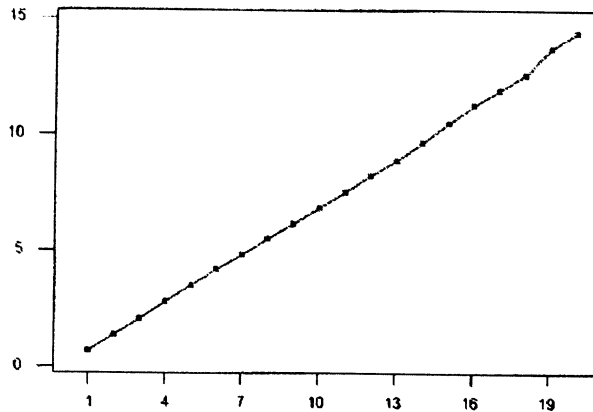


Figura 4.4: O declive do Integral de Correlação em função do número de dimensões para os dados ao dactilografar.

quadrática foi removida antes da análise. Porque os dados tinham ruído, foi feita uma análise *SVD*-Decomposição de um valor singular, isto é, foi utilizado um método para a redução de ruído.

Um gráfico de recorrência destes dados é calculado com um atraso de 9 e a dimensão de mergulho 13 porque a dimensão fractal (cerca de 6.0) era ainda muito elevada após redução do ruído. O raio mínimo para os pontos de recorrência, r , é 1.0 e a linha é definida por três pontos sucessivos na diagonal.

Apesar da função de autocorrelação não decresça rapidamente, o tempo de correlação é de aproximadamente 4 atrasos. Porque a função de informação mútua tem o primeiro mínimo perto de 7, a seguinte análise foi feita com atraso temporal 7. Era necessário pelo menos um mergulho de dimensão 7 porque os falsos vizinhos ocorrem em dimensões mais baixas.

Quando os declives dos integrais de correlação são calculados para valores de $\log(r)$ entre -2.0 e -1.5 , a figura 4.5 mostra que o declive não estabiliza com o aumento da

dimensão de mergulho. Sem redução de ruído, o determinismo não linear não é evidente para estes dados.

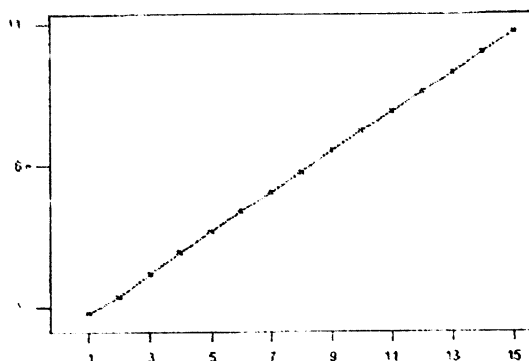


Figura 4.5: O declive do Integral de Correlação em função do número de dimensões de mergulho para os dados da taxa cardíaca.

De seguida, vamos analisar dados de um EEG de um gato, as 5624 observações de um EEG, usando integrais de correlação.

Na figura 4.6 há dois picos discerníveis, porque a função de informação mútua tem o seu primeiro mínimo com o atraso temporal de 2. A máxima dimensão de mergulho é de 20.

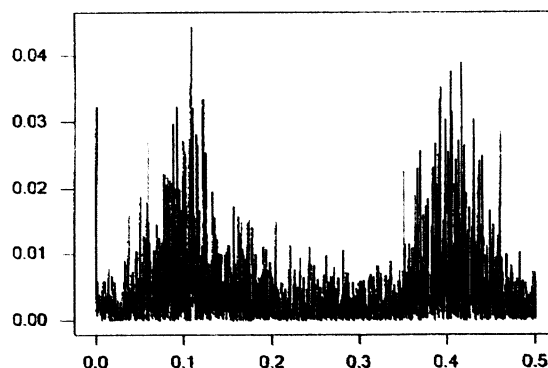


Figura 4.6: Força da frequência em unidades arbitrárias.

Sem procedimentos apropriados para a redução de ruído, não há evidência de determinismo não linear. A figura 4.7 mostra os integrais de correlação.

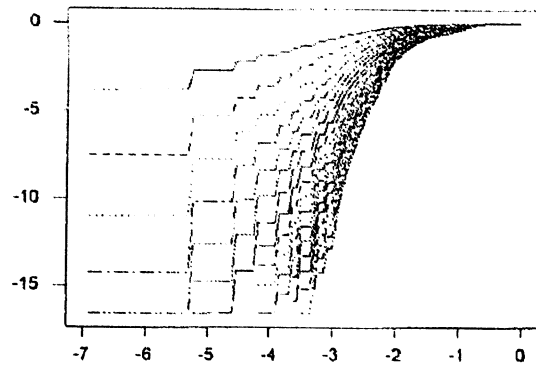


Figura 4.7: Integral de Correlação para os dados de um EEG.

A figura 4.8 mostra a relação crescente entre o declive do integral de correlação e a dimensão de mergulho. O declive é calculado ao longo do intervalo de $\log(r)$ entre -3.1 e -2.5 .

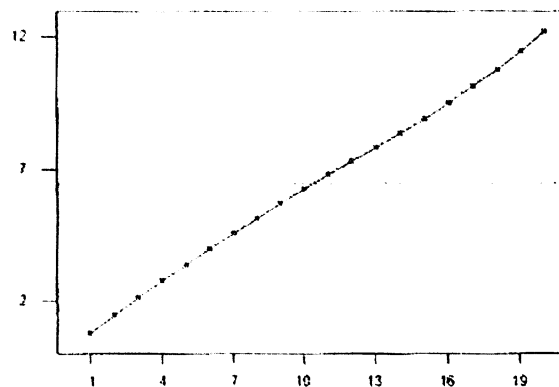


Figura 4.8: Integrais de correlação em função da dimensão de mergulho.

Um declive ilimitado sugere que o EEG tem ruído.

Vamos em seguida expor algumas aplicações dos Expoentes de Lyapunov na Psicologia.

As séries temporais são dominadas pelo ruído, e as estimativas da dimensão fractal podem ser calculadas, no entanto não são de confiança.

Por conseguinte, é pouco útil calcular expoentes Lyapunov para estes conjuntos de dados. No entanto, para comparação com cálculos semelhantes obtidos com redução do ruído, o máximo expoente de Lyapunov foi calculado para a digitação IKT e EEG do gato introduzidos atrás. Para efeitos de comparação o máximo expoente Lyapunov também foi calculado para 2000 observações da série do atrator de Hénon.

Estas análises das séries temporais experimentais são feitas num mergulho de dimensão 10 e a série de Hénon com dimensão 2.

Dado que o número de observações em cada série temporal é bastante pequeno, o maior Expoente de Lyapunov é calculado com a técnica de Rosenstein et al..

Para séries temporais experimentais, houve um mergulho de dimensão 10, atrasos de 5 e 2 e o número de iterações foi de 50.

Cada análise prevê a distância de pontos vizinhos no atrator no tempo t . O declive deste gráfico antes da sua assíptota fornece uma estimativa do maior Expoente de Lyapunov.

Como se vê na figura 4.9 o declive é 0.41 para o atrator de Hénon, um valor próximo de $\lambda_1(0.445)$ presente na tabela 3.11.

As figuras 4.10 e 4.11 mostram os resultados correspondentes dos dados adquiridos ao digitar uma informação numa máquina de escrever e dos dados EEG de um gato, respectivamente. Em ambos os casos nos gráficos há dois segmentos lineares antes da assíptota, tornando difícil o cálculo do Expoente Máximo de Lyapunov.

Fell, Röschke e Beckmann em 1993 observaram evidência para não linearidade nos

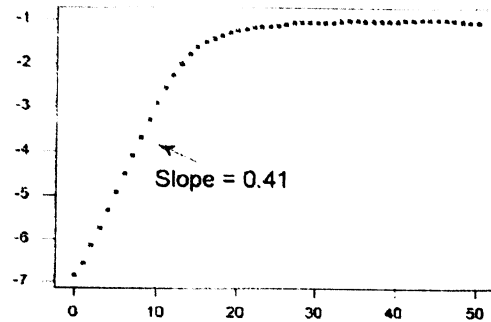


Figura 4.9: Cálculo do máximo expoente de Lyapunov para a série temporal da função de Hénon usando o método de Rosentein et al.

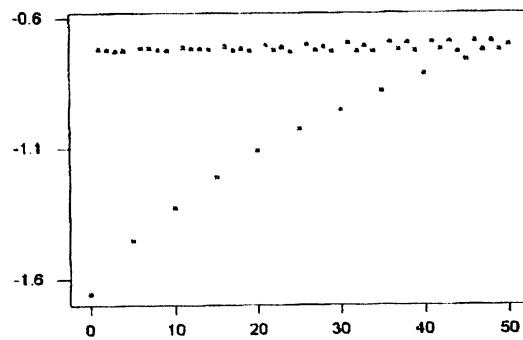


Figura 4.10: Cálculo do máximo expoente de Lyapunov para os dados ao dactilografar usando o método de Rosentein et al.

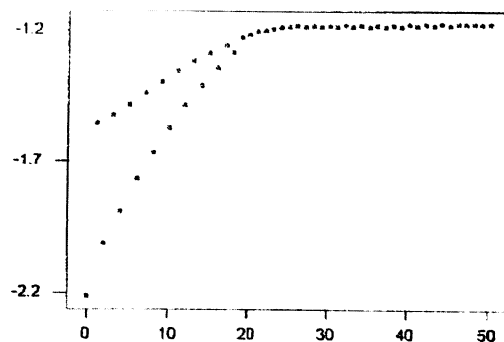


Figura 4.11: Cálculo do máximo expoente de Lyapunov para a série temporal de um EEG usando o método de Rosentein et al.

registos de um EEG durante vários estágios de sono. A estimativa para o maior expoente de Lyapunov obtido usando o algoritmo de Wolf depende do valor do parâmetro seleccionado para o desenvolvimento dos cálculos. Numa análise de cerca de 16 000 observações de um EEG mostrou que o maior expoente de Lyapunov decresce com o crescimento da dimensão de mergulho, obtendo estabilidade além de dimensão 10. O tempo de evolução usado para projectar trajectórias próximas deve ser suficientemente grande para gerar expoentes de Lyapunov estáveis. Analogamente, é também necessário um valor suficientemente grande para a distância máxima na qual a trajectória do atractor se desenvolve no espaço de fases. Os parâmetros recomendados para o uso do algoritmo de Wolf são 40 tempos e dimensão de mergulho 10. Mediram EEG em 4 localizações, P_z , C_z , C_3 e C_4 durante durante vários estágios de sono. Para séries temporais (16384 pontos) registados de C_z , o maior expoente de Lyapunov, variou entre 2.4 a 3.4, diminuiu com o aumento da profundidade do sono. No entanto, não há diferença entre expoentes para estádios 3 e 4. Interpretou que o maior expoente de Lyapunov como um parâmetro de processamento de informação flexível sendo mais baixo durante estádios de sono mais profundos. Numa pesquisa, três anos mais tarde, mostrou-se que um EEG com sono é não linear e provavelmente determinístico. No entanto, não houve evidência directa de existência de caos.

Há um grande número de estudos sobre dinâmicas não lineares em EEG e registos de taxas cardíacas em indivíduos saudáveis e com alguma doença. Na maioria dos casos há uma evidência clara para não linearidade. No entanto, a possibilidade das séries temporais serem de baixa dimensão e caóticas é difícil de discernir. Muito poucos estudos relacionaram os índices não lineares de várias zonas da cabeça aos vários tipos funções do cérebro e comportamento. No entanto, tais índices representam pistas de diagnóstico

úteis que auxiliam o trabalho de neurologistas e neuropsicólogos clínicos, analogamente as análises não-lineares de séries de frequência cardíaca estão a começar a revelar subtis precursores de diagnóstico de uma iminente insuficiência cardíaca.

Segundo [7], Middleton e seus colegas trabalharam no sentido de fazer analogias interessantes relacionando tipos de atrator e tipos de personalidade. Eles descreveram quatro tipos de atratores geralmente dentro destes tipos: ponto fixo, ciclo-limite/periódico, quase periódico/tórus e atratores estranhos. Referiram o seguinte: "Se a personalidade actua como ponto-fixo logo individualmente esperamos que tendam sempre para um comportamento singular. Sem influências exteriores ou mudança de condições iniciais o comportamento emerge sempre e permanece daí em diante, é talvez possível imaginar desordem da personalidade Schizoid ou Esquizofrenia Catatonic gerando este tipo de atrator, o ponto fixo.

No entanto, os mesmos argumentaram que a personalidade não é um ponto fixo, periódico nem um atrator quase periódico: " Se a personalidade é governada por um atrator quase periódico, o comportamento muda ao longo do tempo e esta mudança deverá ocorrer com uma regularidade exacta, onde não há mudanças drásticas."

Concluíram que a personalidade é um atrator estranho, e aparece apresentada por um modelo interaccionista na qual conta com um comportamento limitado e flutuações dinâmicas que ocorrem com pressões internas e externas.

Argumentaram ainda que: "O comportamento exacto será imprevisível de um momento para o outro mas manter-se-á dentro de uma vizinhança do atrator estranho; nenhum comportamento tem igual probabilidade de ocorrência". Em contraste, se o comportamento é aleatório então todo o comportamento possível tem a mesma probabilidade

de ocorrência em qualquer tempo. Os comportamentos complexos requerem modelos que sejam capazes de não serem apenas estáveis e não apenas instáveis. Até os sistemas dinâmicos serem descobertos, os cientistas que estudavam a personalidade e tinham dificuldade em tornar o tempo uma posição interaccionista com significado. Era a teoria que descrevia estabilidade e alguns deles instabilidade, mas não há uma teoria que tenham ambos.

O caos e, mais recentemente, a teoria da Complexidade forneceram-nos a habilidade de compreender ambos os lados. A analogia da personalidade com um atractor estranho, com as suas conotações de parâmetros limitados e flutuações dinâmicas sensíveis parecem fazer sentido.

Hannah, Middleton e Dibello assim como Heiby, também notaram a necessidade de abordagem idiográfica no desempenho de participantes singulares, estudos de séries temporais pois notaram que as complexidades envolvidas com o uso da teoria do Caos dificultam o trabalho do pesquisador.

Hannah faz uso do método dos atrasos de Takens, reconstruiu atractores estranhos, estimou dimensões fractais e os expoentes de Lyapunov, e encontrou que entre flutuação com ruído os temperamentos estendem-se numa ordem considerável. Por outras palavras, caos de baixa dimensão.

Capítulo 5

Conclusão

Em dinâmica procura-se entender a evolução de sistemas cujo estado se altera com o decorrer do tempo, de modo a prever em que estado o sistema se encontrará em dado momento no futuro. A evolução de muitos dos sistemas que encontramos na natureza, mesmo regidos por leis de evolução simples, é geralmente caótica.

Os sistemas caóticos, que inicialmente se supunha ser de uma abordagem quase impossível, inacessíveis à compreensão matemática, têm gradualmente vindo a ser estudados com cada vez maior eficácia, com a ajuda de técnicas e teorias matemáticas ajustadas a esses tipos de sistemas, que forçosamente não se enquadram nas teorias mais clássicas.

A aplicação de técnicas de dinâmicas não lineares para analisar dados de uma grande variedade de aplicações relevantes da psicologia, permite-nos interpretar fenómenos comportamentais com uma nova abordagem. Um grande número de fenómenos, desde funções neurais básicas, registos fisiológicos com significado psicológico, a processos de conhecimento e desenvolvimento, exibem estruturas dinâmicas subjacentes semelhantes. Isto leva-nos a uma nova unificação dos conceitos da ciência comportamental.

Dado que muitos fenómenos de interesse para os psicólogos parecem ter raiz num processo natural, alguns exemplos incluem aprendizagem, emoções, cognição, percepção,

desenvolvimentos fisiológicos e psicoterapia. O cálculo de médias, desvios-padrões, t-tests, qui-quadrado, ANOVA e procedimentos multivariados e correlações Spearman-Roe são práticas comuns que serão melhor aplicadas em situações estáticas. No mínimo, estes procedimentos beneficiariam ser complementados com a análise do caos.

A análise de séries temporais não lineares é um tema muito rico quer por proporcionar a aplicação directa de conceitos, propriedades e técnicas de Matemática, quer por permitir a interface com outras áreas do conhecimento.

O comportamento caótico representa uma das características mais importantes dos sistemas não lineares. A propriedade fundamental de um sistema caótico é a sensibilidade à condição inicial, isto é, condições iniciais muito próximas geram trajectórias inicialmente próximas que, após um certo tempo, passam a apresentar comportamentos completamente não correlacionados entre si. Daí a possibilidade de se fazer previsões a longo prazo baseados em modelos matemáticos quando se tem um sistema caótico.

Para identificarmos e caracterizarmos o comportamento caótico é necessário fazer uso de várias técnicas diferentes, tais como diagrama de bifurcação, expoentes de Lyapunov e valores de entropia.

Os modelos matemáticos são muito úteis, mas muitas vezes é necessário analisar sistemas sem que se conheçam pormenores sobre a sua dinâmica, como no caso de dados obtidos experimentalmente. Assim, nem sempre é possível criar um modelo simplificado que represente o problema, sendo necessário termos outras técnicas como o da reconstrução do espaço de estados.

A base para a maioria dos métodos de análise de séries temporais com dinâmicas não lineares é utilizar a técnica de mergulho e atrasos temporais. A primeira aplicação desta

técnica a séries temporais caóticas, segundo [28], foi feita por Glass and Mackey em 1979, no contexto de equações diferenciais, e foi generalizada em 1980, por Packard baseada numa sugestão de Ruelle. A determinação da dimensão de mergulho e o atraso são de extrema importância e vários métodos podem ser usados para a sua determinação.

A abordagem dos sistemas caóticos representa uma mudança significativa dos métodos tradicionais usados na pesquisa psicológica.

Estudamos diferentes maneiras de quantificar um sistema caótico, procurando de alguma maneira quantitativa especificar o "quanto" o sistema é caótico.

Outra ferramenta na análise de séries temporais é a medida de entropia. A medida de entropia de uma série temporal é importante para a caracterização do sistema dinâmico, indicando qual a quantidade de desorganização, assim a entropia nula ocorre quando o sistema não apresenta um comportamento caótico.

Os expoentes de Lyapunov e a entropia de Kolmogorov-Sinai focam o comportamento do sistema em função do tempo. Se o sistema tem pelo menos um expoente de Lyapunov positivo, podemos dizer que o comportamento do sistema é caótico.

Concluo referindo que nenhum índice quantitativo em particular é o melhor para caracterizar um sistema não linear. Todos os índices quantitativos propostos requerem vários cálculos computacionais para extraírem os seus valores numéricos através de dados. É necessária ter uma atenção especial para ultrapassar dificuldades na sua avaliação. Uma vantagem dada aos expoentes de Lyapunov é que captam a essência das dinâmicas caóticas: a divergência rápida de trajectórias próximas umas das outras. E de que os expoentes de Lyapunov para um espaço de fases multidimensional pode ser relacionado com a dimensão fractal.

A riqueza dinâmica de muitos fenómenos caóticos concentra-se em regiões para onde convergem os estados futuros do sistema, os atractores, possuindo formas geométricas que podem ser bastante complicadas. A estrutura desses atractores ditos estranhos, que numa primeira abordagem poderá parecer completamente irregular e desprovida de qualquer padrão, revela-se, muitas vezes, similar em diferentes escalas e com alguma regularidade na forma como essa regularidade vai aparecendo nas diversas escalas.

A aparente irregularidade desses atractores está relacionada com falta de enquadramento da sua dimensão nos "padrões tradicionais" de dimensão inteira, em que as linhas possuem dimensão um, planos possuem dimensão dois... . Contudo, é possível associar a muitos desses objectos um número fraccionário que generaliza para esses objectos (por isso dito fractais) o conceito de dimensão.

O estudo de dinâmicas não lineares tem sido aplicado em todos os campos da ciência, ele ajuda-nos a descrever, organizar e quantificar o comportamento complexo.

Existem na literatura muitos artigos em que algumas das técnicas estudadas são aplicadas. Temas de investigação muito recentes, na área da psicologia, entre muitos outros, usam invariavelmente algumas das técnicas estudadas neste trabalho.

Acresce ainda dizer que todas as técnicas estudadas têm alguma particularidade que torna cada uma especial do ponto de vista matemático.

Há três princípios básicos da teoria do Caos que têm relevância para a psicologia clínica.

A mudança não é necessariamente linear, isto é, pequenas causas podem ter efeitos grandes.

Determinismo e predicabilidade não são sinónimos, equações determinísticas podem trazer resultados imprevisíveis, caóticas.

Em sistemas que estão "longe do equilíbrio", isto é, caóticos, a mudança não está relacionada com causas externas. Tais sistemas auto organizam-se num nível superior de organização.

É este contacto com as experiências que nos leva a acreditar que desenvolvimentos crescentes em dinâmica não linear nos tragam uma contribuição permanente para o campo do mundo científico.

Dinâmicas não lineares e caos enriqueceram o nosso vocabulário tanto verbal como conceitual e deu-nos novos modelos que nos ajudam a compreender o mundo que nos rodeia.

Bibliografia

- [1] ADDISON, Paul S.. *Fractals and Chaos-An illustrated course*. Edinburgh: Institute of Physics Publishing, 1997.
- [2] ALLIGOOD, Kathleen T., SAUER, Tim D., YORKE, James A, *Chaos-An introduction to dynamical systems*. New York: Springer-Verlag New York, Inc., 1996.
- [3] BASK, M., GENÇAY, R.. *Testing chaotic dynamics via Lyapunov exponents*. Physica D: Nonlinear Phenomena, Volume 114, 1-2, 15 March 1998, 1-2.
- [4] BROCKWELL, Peter J., DAVIS, Richard A.. *Introduction to Time Series and Forecasting*. New York: Springer-Verlag, 1996.
- [5] BOX, George E., Jenkins, Gwilym M.. *Time Series analysis forecasting and control*. California: Holden-Day Inc., 1976.
- [6] BAKER, G. L., GOLLUB, J. P.. *Chaotic dynamics-An introduction*. Canada: Cambridge University Press, 1990.
- [7] BERRY, Michael V., PERCIVAL, I.C., WEISS, N.O.. *Dynamical Chaos*. New Jersey: Princeton University Press, 1987.

-
- [8] BROWN, R. BRYANT, P., ABARBANEL, H. D.I.. *Computing the Lyapunov spectrum of a dynamical system from an observed time series*. Physical Review A, volume 43, 1991, 2787–2806.
- [9] BÜTZ, Michael R., *Chaos and Complexity. Implications for psychological Theory and Practice*. Wyoming: Taylor&Francis, 1997.
- [10] CACIOPPO, John T., TASSINARY, Louis G.. *Principles of Psychophysiology-Physical, social, and inferential elements*. New York: Cambridge University Press, 1990.
- [11] CHATFIELD, Chris. *Time-Series Forecasting*. Florida: Chapman & Hall/CRC, 2000.
- [12] CRYER, Jonathan D.. *Time Series Analysis*. Boston: PWS Publishers, 1986.
- [13] DEVANEY, Robert L. *An Introduction to Chaotic Dynamical Systems*. Second Edition. Addison-Wesley Publishing Company. Boston, 1988.
- [14] GALLEZ, D., BABLOYANTZ, A.. *Predictability of a human EEG: a dynamic approach*. Biological Cybernetics, Volume 64, 5 March 1991, 381-391.
- [15] GENÇAY, R.. *Nonlinear prediction of noisy time series with feedforward networks*. Physics Letters A, volume 187, 5-6, 2 May 1994, 397-403.
- [16] GENÇAY, R.. *A statistical framework for testing chaotic dynamics via Lyapunov exponents*. Physica D: Nonlinear Phenomena, volume 89, 3-4, 1 January 1996, 261-266.

-
- [17] GENÇAY, R., DECHERT, W. D.. *An algorithm for the n Lyapunov exponents of an n -dimensional unknown dynamical system*. Physica D: Nonlinear Phenomena, Volume 59, 1-3, 142-157.
- [18] GONÇALVES, Esmeralda, LOPES, Nazaré Mendes. *Séries Temporais. Modelações lineares e não lineares*. Sociedade Portuguesa de Estatística, 2008.
- [19] GOLIA, S., SANDRI, M.. *Resampling chaotic time series*. Physical Review Letters, volume 78, number 22, 2 June 1997.
- [20] GUASTELLO, Stephen J., *Chaos, Catastrophe, and human Affairs. Applications of Nonlinear Dynamics to work, Organizations, and Social Evolution*. Mahwah, New Jersey: Lawrence Erlbaum Associates, Publishers, 1995.
- [21] HAMILTON, James D.. *Time Series Analysis*. New Jersey: Princeton University Press, 1994.
- [22] HEATH, Richard A.. *Nonlinear Dynamics-Techniques and Applications in Psychology*. Mahwah: Lawrence Erlbaum Associates, Inc., 2000.
- [23] HILBORN, Robert C.. *Chaos and Nonlinear Dynamics an introduction for scientists and engineers*. New York: Oxford University Press, Inc., 1994.
- [24] HOLDEN, Arun V.. *Chaos*. Manchester:Manchester University Press, 1996.
- [25] HOLMGREN, Richard A.. *A First Course in Discrete Dynamics*. Second Edition. Springer, 1996.

-
- [26] KANTZ, H.. *A robust method to estimate the maximal Lyapunov exponent of a time series*. Physics Letters A, Volume 185, 1, 31 January 1994, 77-87.
- [27] KAPITANIAK, Tomasz. *Chaos in Systems with Noise*. Singapore: Word Scientific Publishing Co. Pte. Ltd., 1990.
- [28] KAPLAN, Daniel, GLASS, Leon. *Understanding Nonlinear Dynamics*. New York: Springer-Verlag, 1995.
- [29] KENDALL, Sir Maurice, ORD, J. Keith. *Time Series*. Great Britain: Edward Arnold, 1990.
- [30] LYNCH, Stephen. *Dynamical Systems with applications using MAPLE*. Boston: Birkhäuser, 2001.
- [31] LORENZ, Edward. *The Essence of Chaos*. London: UCL Press Limited ,1993
- [32] MONAGEM, M.B., GEDDES, K.O., HEAL, K.M., LABAHN,G., VORKOETTER, S.M.. *MAPLE V- Programming Guide*. New York: Springer-Verlag, 1998.
- [33] MOON, Francis C.. *Chaotic and Fractal Dynamics-An introduction for Applied Scientists and Engineers*. New York: John Wiley & Sons, Inc., 1992.
- [34] MORY, H.. *Fractal dimensions of chaotic flows of autonomous dissipative systems*. Progress in Theoretical Physics, 1980.
- [35] MURTEIRA, J. F. Bento, MÜLLER, Daniel A.. TURKMAN, K. Feridun. *Análise de Sucessões Cronológicas*, Mc Graw Hill, 1993.

-
- [36] PRIESTLEY, M. B.. *Non-linear and Non-Stationary time series analysis*. Great Britain: Academic Press Limited, 1991.
- [37] PEITGEN, Heinz-Otto, JÜRGENS, Hartmut, SAUPE, Diemar. *Chaos and fractals*. New Frontiers of Science. New York: Springer-Verlag, 1992.
- [38] PEITGEN, Heinz-Otto, JÜRGENS, Hartmut, SAUPE, Diemar. *Fractals for the Classroom*. Vols.1 e 2. New Frontiers of Science. New Yorke: Springer-Verlag, 1992.
- [39] RAO, T. Subba. *Time Series analysis*. London: Chapman & Hall, 1993.
- [40] ROBERTSON, Robin, COMBS, Allan. *Chaos Theory in Psychology and the Life sciences*. New Jersey: Lawrence Erlbaum Associates, Publishers, 1995.
- [41] ROBINSON, Clark . *Dynamical Systems – Stability, Symbolic Dynamics and Chaos*. Second Edition. Boca Raton: CRC Press, 1999.
- [42] ROSENSTEIN, M. T., COLLINS, J. J. & DE LUCA, C. J.. *A practical method for calculating largest Lyaponov from small data sets*. Physica D: Nonlinear Phenomena Volume 65, 1-2, 15 May 1993, 117-134.
- [43] SCHREIBER, T.. *Interdisciplinary application of nonlinear time series methods*. Physics Reports, Volume 308, 1 January 1999, 1-64.
- [44] SCHUSTER, Heinz Georg. *Deterministic Chaos – an Introduction*. Weinheim: VCH Verlagsgesellschaft mbH, 1995.
- [45] STROGATZ, Steven H.. *Nonlinear Dynamics and Chaos*. Cambridge: Perseus Books Publishing. L.L.C., 1994.



- [46] WOLF, A., SWIFT, J. B., *Determining Lyapunov exponents from time series*. Physica D: Nonlinear Phenomena, Volume 16, 3, July 1985, 285-317.
- [47] <http://sprott.physics.wisc.edu/chaos/lyapexp.htm>
- [48] <http://info.lboro.ac.uk/departments/ma/gallery>
- [49] <http://www.cienciahoje.pt>