

UNIVERSIDADE DE BRASÍLIA
INSTITUTO DE FÍSICA

JOÃO PEDRO VALERIANO MIRANDA

**ESTUDO DO MAPA LINEAR N-DIMENSIONAL COM
RUÍDO ADITIVO PELO MÉTODO DA FUNÇÃO
CARACTERÍSTICA DE PAUL LÉVY**

BRASÍLIA

11 DE DEZEMBRO DE 2020

João Pedro Valeriano Miranda

**Estudo do Mapa Linear N-Dimensional com Ruído Aditivo
Pelo Método da Função Característica de Paul Lévy**

Monografia apresentada ao Instituto de Física da Universidade de Brasília como parte dos requisitos necessários à obtenção do título de Bacharel em Física.

Orientador: Annibal Dias de Figueiredo Neto

Universidade de Brasília – UnB

Instituto de Física

Brasília

11 de dezembro de 2020

João Pedro Valeriano Miranda

Estudo do Mapa Linear N-Dimensional com Ruído Aditivo Pelo Método da Função Caraterística de Paul Lévy/ João Pedro Valeriano Miranda . – Brasília, 11 de dezembro de 2020-

92 p. : il. (algumas color.) ; 30 cm.

Orientador: Annibal Dias de Figueiredo Neto

Monografia – Universidade de Brasília – UnB
Instituto de Física, 11 de dezembro de 2020.

1. Processos estocásticos lineares de tempo discreto. 2. Função Característica de Lévy. I. Annibal Dias de Figueiredo Neto. II. Universidade de Brasília. III. Instituto de Física. IV. Estudo do Mapa Linear N-Dimensional com Ruído Aditivo Pelo Método da Função Caraterística de Paul Lévy

CDU 02:141:005.7

João Pedro Valeriano Miranda

Estudo do Mapa Linear N-Dimensional com Ruído Aditivo Pelo Método da Função Característica de Paul Lévy

Monografia apresentada ao Instituto de Física da Universidade de Brasília como parte dos requisitos necessários à obtenção do título de Bacharel em Física.

Brasília, 11 de dezembro de 2020:

Annibal Dias de Figueiredo Neto
Orientador

Olavo Leopoldino da Silva Filho
Convidado 1

Bernardo de Assunção Mello
Convidado 2

Brasília
11 de dezembro de 2020

Dedicado à memória do Professor Ivan Soares Ferreira.

Agradecimentos

Este trabalho simboliza a conclusão de uma fase muito importante da minha vida, definitivamente a melhor, até agora. Tendo isso em mente, me sinto obrigado a agradecer a todos que fizeram parte dessa trajetória.

Em primeiro lugar, agradeço aos meus pais, Rogério e Renata, assim como ao meu irmão, Paulo, pelo amor, pela compreensão e pelo apoio constante e incondicional.

Agradeço, é claro, ao Professor Annibal, orientador desse trabalho, por me guiar, me dar as ferramentas básicas para dar o pontapé inicial, e sanar as dúvidas que surgiram ao longo do caminho. Também preciso agradecer a alguns dos professores que marcaram minha formação, como Júnio, Vanessa, Clóvis, Arsen, Pinzul, Zapata, Pedroza e Viktor, além do Alexandre e do Lucas, que me orientaram no meu período no LNLS. São necessários ainda dois agradecimentos muito especiais. Primeiro, ao Professor Olavo, que me introduziu à pesquisa em física teórica e me proporcionou discussões muito frutíferas, que contribuíram para o desenvolvimento do meu pensamento científico. Segundo, ao Professor Ivan, que me apresentou a pesquisa em física experimental, o que me ensinou a enxergar a física de outra maneira. Se hoje sou um físico preparado para seguir carreira acadêmica, devo isso a vocês.

Preciso ainda agradecer aos meus companheiros de jornada, amigos que conheci através da física, e que hoje, pra mim, são uma família. Primeiro, aqueles que entraram no curso junto comigo: Caio, Graci, Alice, Lettieri, Deborah, Paulo e Vítor Machado; além daqueles que se agregaram com o passar do tempo: João Augusto, Vítor, Tábata, Ranier, Felipe, Lud, Igor, Vinícius, Marcos, Miguel e Luísa; e daqueles que conheci no CBPF: Linda, Milena, Lívia e João Vítor; e no CNPEM: Gabriel e Letícia. Um agradecimento especial ao Pedro, que me acolheu já no dia da minha matrícula e rapidamente se tornou um dos meus melhores amigos. Obrigado por cuidar de mim, e constantemente tentar me ensinar a não me estressar tanto com os trabalhos, apesar de eu sempre falhar. Outro agradecimento especial à Mari que, se antes era “apenas” outra grande amiga que também entraria nessa lista de outra forma, hoje é, na verdade, uma das minhas partes favoritas da vida. Muito obrigado pelo apoio constante, pelo carinho, pela compreensão e pela nossa amizade, que me faz muito feliz. Ainda agradeço aos amigos de fora da física, com quem não pude conviver tanto quanto gostaria, mas que também são muito importantes para mim: Sarah, Letícias, Júlia, Nathália, Miguel e, especialmente, o Léo, a quem agradeço por todos os momentos juntos, por me ensinar tanto sobre tudo, e me inspirar a estudar mais. Vocês todos me ensinaram muito sobre a vida, de diferentes maneiras, e eu me sinto muito sortudo de conhecê-los. Vocês são muito mais importantes para mim do que podem imaginar, e eu sinto que poderia fazer de tudo por vocês.

Resumo

Neste trabalho, estudamos mapas multilineares com ruído aditivo, ou seja, equações de diferença estocásticas lineares, inspirados na discretização de equações diferenciais estocásticas lineares, ao supormos que o ruído tem valor constante por intervalos. Integrando a equação diferencial estocástica em cada intervalo, obtemos uma equação de diferenças estocástica. Mostramos que o Teorema do Limite Central pode ser visto como apenas um caso especial de um teorema descrevendo a distribuição limite de um processo estocástico, e determinamos condições sobre os mapas para que a distribuição limite de cada uma de suas componentes seja gaussiana ou não. Por fim, estudamos a equação de Langevin com ruído constante por intervalos, e mostramos que a distribuição limite de velocidades do sistema não necessariamente será gaussiana, mas conferimos que, reduzindo os intervalos, de modo a retornar à equação diferencial estocástica, recuperamos a distribuição normal. Além disso, consideramos o oscilador harmônico, amortecido ou não, com ruído aditivo, e suas propriedades mais interessantes.

Palavras-chaves: Processos estocásticos de tempo discreto. Sistemas lineares. Função característica. Soma de variáveis aleatórias. Gaussianização.

Abstract

We study multilinear maps with additive noise, that is, stochastic linear difference equations, inspired by the discretization of stochastic differential equations, when we suppose the noise is constant by intervals. Integrating the stochastic differential equation on each interval, we obtain a stochastic difference equation. We show that the Central Limit Theorem can be seen as only a special case of a more general theorem describing the limit distribution of a stochastic process, and determine the necessary conditions for the limit distribution of each component of the system to be gaussian. Last but not least, we study the Langevin equation with the random force constant by intervals, and show that the velocity's limit distribution is not necessarily gaussian but, if we reduce the time intervals to recover the stochastic differential equation, we obtain a normal distribution. Besides that, we consider the harmonic oscillator, damped or not, and its most interesting properties.

Key-words: Discrete time stochastic processes. Linear systems. Characteristic function. Sum of random variables. Gaussianization

Lista de ilustrações

Figura 1 – Ruído constante por intervalos.	18
Figura 2 – Simulação da equação de Langevin para $b = 0.005$, $m = \Delta t = 1$, considerando 10000 trajetórias de 1000 passos.	77
Figura 3 – Simulação da equação de Langevin para $b = 2$, $m = \Delta t = 1$, considerando 10000 trajetórias de 1000 passos.	78
Figura 4 – Oscilador Harmônico num fluido. $b = 5$, $k = m = 1$: Regime supercrítico. $\Delta t = 10$	82
Figura 5 – Oscilador Harmônico num fluido. $b = k = m = 1$: Regime subcrítico. $\Delta t = 0.01$	83
Figura 6 – Oscilador Harmônico com força aleatória, sem amortecimento. $k = m = 1$, $\Delta t = 0.01$	84

Sumário

	Introdução	17
1	MÉTODOS MATEMÁTICOS	23
1.1	Variáveis Aleatórias	23
1.1.1	A Função Característica de uma Variável Aleatória e sua Forma Canônica para Variância Finita	24
1.2	Vetores Aleatórios	28
1.2.1	A Forma Canônica da Função Característica de um Vetor Aleatório com Componentes de Variância Finita	29
1.2.2	O Vetor Principal de um Vetor Aleatório	31
1.2.3	Vetor Principal de Componentes Independentes	32
1.3	A Função Característica de Paul Lévy Para Somas de Variáveis Aleatórias	33
1.3.1	Generalizando o Teorema do Limite Central de Paul Lévy	35
1.4	Expansão em Série de Potências da Função Característica	37
1.4.1	Variáveis Reduzidas Identicamente Distribuídas	37
1.5	Potências de Matrizes	39
1.5.1	A Forma Canônica de Jordan Para Matrizes Não Diagonalizáveis	40
2	O CASO UNIDIMENSIONAL	43
2.1	Ruído Estacionário	44
3	SOLUÇÕES GERAIS PARA A FUNÇÃO CARACTERÍSTICA DO MAPA	47
3.1	Solução na Base do Vetor Principal para o Ruído	48
3.2	Solução na Base que Diagonaliza o Mapa	51
3.3	Solução para o Mapa Não Diagonalizável: Forma Canônica de Jordan	55
4	GAUSSIANIZAÇÃO COM RUÍDO INDEPENDENTE	59
4.1	Matriz diagonalizável com autovalores reais	61
4.2	Matriz diagonalizável com autovalores complexos	64
4.3	Matriz não Diagonalizável: Forma Canônica de Jordan	66
5	EXEMPLOS	73
5.1	A Equação de Langevin	73
5.2	O Oscilador Harmônico num Fluido	78

Conclusão	85
REFERÊNCIAS	87
APÊNDICES	89
APÊNDICE A – TEOREMAS E DEMONSTRAÇÕES	91

Introdução

Dada sua utilidade para modelar fenômenos envolvendo incertezas ou ruídos, processos estocásticos podem ser encontrados em inúmeras áreas de pesquisa, passando por modelos de ecologia ([MARTINEZ-GARCIA et al., 2020](#)), economia ([BLACK; SCHOLLES, 1973](#)), até climatologia ([PALMER, 2019](#)) e terremotos ([SHINOZUKA; DEODATIS, 1988](#)). Frequentemente estes processos estocásticos aparecem em duas formas mais comuns: a primeira delas é através de equações diferenciais parciais que descrevem a evolução espaço-temporal de uma distribuição de probabilidades. O exemplo mais comum é a equação de Fokker-Planck. Outra possibilidade é um sistema de equações diferenciais estocásticas, descrevendo a evolução temporal de um estado do sistema sobre o espaço de fase. Neste caso, o exemplo mais conhecido é a equação de Langevin.

Neste trabalho, vamos considerar o segundo caso, em que acompanhamos a evolução temporal de um estado inicial, impulsionado por uma equação linear com um ruído aleatório. Porém, para lidar com equações diferenciais estocásticas, vamos ter alguns problemas associados às descontinuidades do ruído, como vamos ver a seguir. Consideremos um caso simples:

$$\frac{dx(t)}{dt} = b(t, x(t)) + \sigma(t, x(t)) \cdot \eta(t), \quad (1)$$

onde t é o tempo, x é a variável de interesse, b e σ são funções bem definidas, e η é um ruído aleatório, com uma distribuição de probabilidades qualquer. Note que $\eta(t)$ não é uma função de t , mas sim uma variável aleatória, com diferentes valores definidos aleatoriamente para diferentes instantes t .

Como discutido por ([ØKSENDAL, 2003](#)), em geral, quando consideramos uma aplicação, desejamos que $\eta(t_1)$ e $\eta(t_2)$ sejam independentes para $t_1 \neq t_2$; que $\eta(t)$ seja estacionário, ou seja, que a distribuição de $\eta(t)$ não varie ao longo do tempo; e que o valor médio de $\eta(t)$ seja nulo. Porém, é claro que um ruído que obedeça a estas condições não pode ter uma trajetória contínua, então como podemos incorporá-lo de maneira razoável a uma equação diferencial? Uma saída é discretizar a equação diferencial, considerando intervalos de tempo finitos, de modo a fazermos um tratamento independente da continuidade do ruído, e então tomarmos de volta o limite infinitesimal. Considerando intervalos de tempo finitos, temos:

$$x_{n+1} - x_n = \Delta t_n [b(t_n, x_n) + \sigma(t_n, x_n) \cdot \eta_n]. \quad (2)$$

Aqui, $\Delta t_n = t_{n+1} - t_n$, e seu valor pode variar em torno de uma média dependendo de n . Continuando, podemos transpor para o n -ésimo x a partir do valor inicial, somando equações do tipo (2) de índices 0 até $n - 1$:

$$x_n = x_0 + \sum_{k=0}^{n-1} \Delta t_k b(t_k, x_k) + \sum_{k=0}^{n-1} \Delta t_k \sigma(t_k, x_k) \cdot \eta_k. \quad (3)$$

Este seria o momento de retomar o limite $\Delta t_k \rightarrow 0 \forall k$, mas não sabemos sequer se esse limite existe. É possível mostrar que o limite de fato existe, caso consideremos $\Delta t_k \eta_k = \Delta V_k = V_{k+1} - V_k$, sendo V_k uma variável aleatória distribuída de acordo com o processo de Wiener (movimento browniano). Mais detalhes podem ser encontrados em (ØKSENDAL, 2003). Não entraremos nessa discussão, uma vez que seguiremos com o estudo dessa equação por outro caminho.

Consideremos que o ruído apresenta variações apenas em intervalos finitos, de modo que $\eta(t) = \eta[n + 1] \forall t, t_n < t \leq t_{n+1}$. Ou seja, o ruído assume um valor fixo durante um intervalo de tempo fixo Δt e então muda de valor, como na Figura 1.

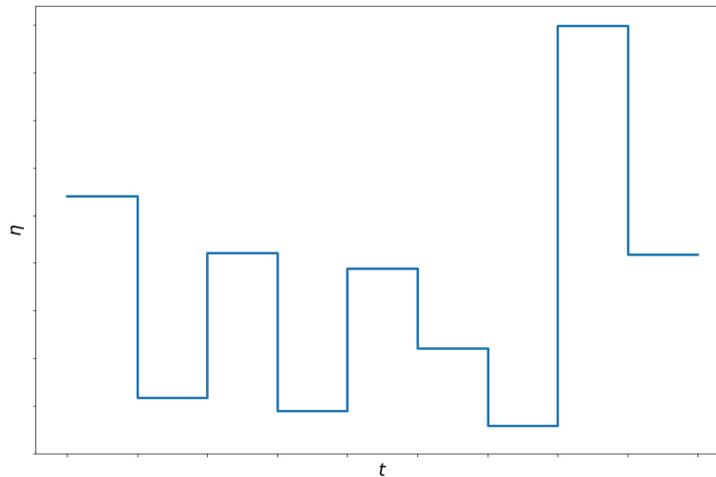


Figura 1 – Ruído constante por intervalos.

Ao longo deste trabalho, vamos considerar apenas modelos lineares ($b(t, x) = Ax$, $\sigma(t, x) = 1$). Podemos integrar a equação entre os momentos t_n e t_{n+1} , escrevendo $\eta[n + 1] = B$:

$$\frac{dx(t)}{dt} = Ax(t) + B. \quad (4)$$

Integrando, obtemos:

$$t_{n+1} - t_n = \Delta t = \int_{x_n}^{x_{n+1}} \frac{dx}{Ax + B} = A^{-1} \ln(Ax + B) \Big|_{x_n}^{x_{n+1}} = A^{-1} \ln \left(\frac{Ax_{n+1} + B}{Ax_n + B} \right) \quad (5)$$

$$\begin{aligned} \Rightarrow Ax_{n+1} + B &= (Ax_n + B) e^{A\Delta t} \\ \therefore x_{n+1} &= e^{A\Delta t} x_n + \frac{B}{A} (e^{A\Delta t} - 1). \end{aligned} \quad (6)$$

Note que A deve ser diferente de zero. Definindo:

$$e^{A\Delta t} \rightarrow A; \quad \frac{e^{A\Delta t} - 1}{A} \rightarrow \sigma, \quad (7)$$

obtemos a equação de diferenças:

$$x[n + 1] = Ax[n] + \sigma\eta[n + 1]. \quad (8)$$

Daqui pra frente, ao longo do texto vamos utilizar essa notação de colchetes para denotar o tempo discreto, aparecendo na frente de variáveis aleatórias.

Desejamos estudar sistemas multilineares, ou seja, com várias componentes:

$$\frac{d\mathbf{x}(t)}{dt} = \mathbb{A}\mathbf{x}(t) + \eta(t), \quad (9)$$

onde \mathbf{x} é então um vetor l -dimensional, assim como η , e \mathbb{A} é uma matriz $l \times l$. Como no caso unidimensional, vamos integrar sobre um intervalo Δt , considerando o ruído constante, agora vetorial, \mathbf{B} . Inspirados no caso unidimensional, esperamos uma solução do tipo

$$\mathbf{x}(t) = e^{\mathbb{A}t} \mathbf{x}_0 + \mathbb{A}^{-1} (e^{\mathbb{A}t} - \mathbb{1}) \mathbf{B}, \quad (10)$$

onde $\mathbb{1}$ é a matriz identidade. Note que \mathbb{A} deve ser inversível. Caso contrário, será necessário discretizar o sistema de outra maneira, como veremos no Capítulo 5. Substituindo na equação

$$\frac{d\mathbf{x}(t)}{dt} = \mathbb{A}\mathbf{x}(t) + \mathbf{B}, \quad (11)$$

verificamos que esta é, de fato, a solução. Portanto, temos:

$$\mathbf{x}[n + 1] = e^{\mathbb{A}\Delta t}\mathbf{x}[n] + \mathbb{A}^{-1}\left(e^{\mathbb{A}\Delta t} - \mathbb{1}\right)\mathbf{B}. \quad (12)$$

Novamente, redefinindo $e^{\mathbb{A}\Delta t} \rightarrow \mathbb{A}$ e $\mathbb{A}^{-1}\left(e^{\mathbb{A}\Delta t} - \mathbb{1}\right) \rightarrow \sigma$, concluímos:

$$\mathbf{x}[n + 1] = \mathbb{A}\mathbf{x}[n] + \sigma\eta[n + 1] \quad (13)$$

Esta equação será o foco deste trabalho. A partir dela, desejamos avaliar a distribuição de \mathbf{x} no limite $n \rightarrow \infty$, em termos das propriedades das matrizes \mathbb{A} e σ . Estaremos lidando apenas com sistemas lineares dada a sua simplicidade. Mesmo assim, os resultados aqui apresentados podem ser interessantes também para sistemas mais complexos, uma vez que em muitos casos, eles são equivalentes às suas versões linearizadas ao redor do ponto de equilíbrio, de acordo com o Teorema de Hartman-Grobman ([HARTMAN, 1963](#)).

Podemos tratar o problema como uma soma de variáveis aleatórias, uma vez que cada novo passo $x[n + 1]$ do mapa (13) depende de uma nova variável aleatória $\eta[n + 1]$, desconsiderando o termo relacionado à condição inicial $x[0]$.

Desejamos analisar o sistema através da função característica de Paul Lévy ([LÉVY, 1924](#)) – que simplifica imensamente a tarefa de avaliar somas de variáveis aleatórias – com o objetivo de determinar as condições necessárias para a gaussianização da distribuição limite do mapa

Com a evolução do sistema (13), é de se esperar que precisaremos calcular potências da matriz \mathbb{A} . Para simplificar a operação, podemos realizar uma mudança de base no sistema, passando para uma base que diagonalize \mathbb{A} , ou passe para sua forma de Jordan, caso \mathbb{A} não seja diagonalizável. Como será visto no Capítulo 3, a mudança de base do sistema nos obriga a eventualmente analisar um o problema da soma de variáveis aleatórias correlacionadas entre si.

Todas as técnicas matemáticas necessárias à realização do trabalho serão detalhadas no Capítulo 1. Caso o leitor já esteja familiarizado com as técnicas básicas apresentadas, ele pode seguir diretamente para o Capítulo 2, no qual estudamos o caso unidimensional, cuja análise é razoavelmente simples, e mostramos como a distribuição limite de um sistema pode não seguir o TLC. Trata-se de uma importante motivação para o trabalho, uma vez que explicita a existência de sistemas que não convergem para distribuições gaussianas, além de apresentarem difusão não normal.

No Capítulo 3 apresentamos o sistema vetorial a ser estudado e apresentamos a forma geral da função característica da distribuição do sistema. Como será visto, há mais de uma maneira de analisar o problema, cada uma com suas vantagens e desvantagens.

No Capítulo 4, avaliamos um caso particular no qual as componentes do ruído não apresentam correlações entre si e mostramos que, neste caso, a gaussianização do sistema é determinada pelo maior módulo entre os autovalores da matriz A .

Por fim, no Capítulo 5 estudamos a equação de Langevin discretizada, e mostramos que a distribuição de velocidades pode não ser gaussiana. Além disso, este é um bom exemplo de como um sistema, por mais simples que seja, pode apresentar uma matriz não inversível, não se adequando à expressão (12) para sua discretização, e exigindo que façamos uma análise mais cuidadosa. De qualquer modo, após a discretização, utilizamos as ferramentas elaboradas ao longo do trabalho para avaliar o comportamento do sistema.

1 Métodos Matemáticos

1.1 Variáveis Aleatórias

Uma variável aleatória pode ser entendida como uma variável cujo valor é resultado de um processo aleatório, que obedece uma determinada função distribuição de probabilidades $f(x)$, ou $f(x, t)$ caso função distribuição varie ao longo do tempo.

Vamos apresentar uma definição um pouco mais formal e útil (TOME; OLIVEIRA, 2015). Podemos definir uma variável aleatória x como uma variável que pode assumir diferentes valores associados a pontos de um espaço Ω , que podemos chamar espaço amostral.

Se Ω é um conjunto discreto como, por exemplo, o conjunto dos números naturais, podemos associar a cada ponto x_i de Ω , que x pode assumir, um número p_i , com a seguinte restrição:

$$\sum_{\{x_i \in \Omega\}} p_i = 1. \quad (1.1)$$

Neste caso, p_i é a probabilidade de o processo aleatório associado à variável aleatória em questão resulte em x_i .

Porém, se Ω é um conjunto contínuo, devemos trocar o somatório por uma integral, e é conveniente definir uma função $\rho(x)$, que chamamos densidade de probabilidade (nome cujo significado ficará claro em breve), de modo que, no caso de uma variável unidimensional, a probabilidade de se obter um valor para x no intervalo $[a, b]$ é dada por:

$$P(a \leq x \leq b) = \int_a^b \rho(x) dx. \quad (1.2)$$

Neste caso, uma restrição análoga a (1.1) é:

$$\int_{-\infty}^{\infty} \rho(x) dx = 1. \quad (1.3)$$

No caso contínuo, não é bem definida a probabilidade de x para um ponto de Ω , mas podemos dizer, baseados na Equação (1.2), que a probabilidade de que a variável x esteja num intervalo entre x_0 e $x_0 + dx$ é $\rho(x_0) dx$.

Com estas definições, obtemos fórmulas simples para calcular os momentos estatísticos de uma variável aleatória. Se queremos calcular o valor médio de uma função da nossa variável aleatória, devemos calcular essa função para cada valor possível da variável, e então tomar uma média ponderada destes resultados, onde o peso é exatamente a probabilidade de cada possível valor de x em Ω . Ou seja, no caso discreto temos:

$$\langle f(x) \rangle = \sum_{\{x_i \in \Omega\}} f(x_i) p_i, \quad (1.4)$$

e no caso contínuo:

$$\langle f(x) \rangle = \int_{-\infty}^{\infty} f(x) \rho(x) dx. \quad (1.5)$$

Os momentos estatísticos μ_i são definidos como os valores médios de potências da variável aleatória, e fornecem informações sobre o formato da distribuição de probabilidades da variável em questão. De acordo com as Equações (1.4) e (1.5), temos:

$$\mu_n = \langle x^n \rangle = \begin{cases} \sum_{\{x_i \in \Omega\}} x_i^n p_i, & \text{se } \Omega \text{ é discreto;} \\ \int_{-\infty}^{\infty} x^n \rho(x) dx, & \text{se } \Omega \text{ é contínuo.} \end{cases} \quad (1.6)$$

Neste trabalho, lidaremos apenas com distribuições contínuas, e consideraremos ruídos independentes entre si, ou seja, não correlacionados. Matematicamente, isto significa dizer que a covariância de dois ruídos quaisquer será nula:

$$\text{Cov}(\eta_i, \eta_j) = \langle (\eta_i - \langle \eta_i \rangle)(\eta_j - \langle \eta_j \rangle) \rangle = m_i^2 \delta_{i,j}, \quad (1.7)$$

sendo $m_i = \langle (\eta_i - \langle \eta_i \rangle)^2 \rangle^{1/2}$ o desvio padrão de η_i .

1.1.1 A Função Característica de uma Variável Aleatória e sua Forma Canônica para Variância Finita

Definimos a função característica de uma variável aleatória x como:

$$\psi_x(z) = \langle e^{Izx} \rangle, \quad (1.8)$$

sendo I a unidade imaginária, $I^2 = -1$. Esta definição é muito útil por permitir o cálculo dos momentos estatísticos da variável aleatória de modo muito simples. Note que a n -ésima derivada da função característica é dada por:

$$\frac{d^n}{dz^n} \psi_x(z) = \langle (Ix)^n e^{Izx} \rangle. \quad (1.9)$$

Tomando $z = 0$:

$$\left. \frac{d^n}{dz^n} \psi_x(z) \right|_{z=0} = \langle (Ix)^n \rangle \quad (1.10)$$

$$\mu_n = \langle x^n \rangle = (-I)^n \left. \frac{d^n}{dz^n} \psi_x(z) \right|_{z=0}. \quad (1.11)$$

Ou seja, obtemos os momentos estatísticos simplesmente derivando a função característica, o que significa que podemos escrever a função característica como uma série de potências dos momentos estatísticos:

$$\psi_x(z) = \sum_{k=0}^{\infty} \frac{\mu_k}{k!} z^k = \sum_{k=0}^{\infty} \frac{\langle x^k \rangle}{k!} (Iz)^k \quad (1.12)$$

Como veremos na seção 1.3, o uso da função característica ainda apresenta outras vantagens que serão muito úteis ao longo deste trabalho. Note que, pela definição (1.11), o momento estatístico de ordem zero, μ_0 , deve ser exatamente a normalização da densidade de probabilidade, ou seja, $\mu_0 = 1$.

Caso a variável aleatória em questão possua variância finita, a função característica pode ser escrita numa forma particularmente conveniente, que chamamos forma canônica da função característica, e será apresentada a seguir.

Pela definição da função característica, se escrevemos a exponencial em série de Taylor, temos:

$$\begin{aligned} \psi_x(z) &= \left\langle 1 + Ixz - \frac{1}{2}(xz)^2 + o(z^2) \right\rangle \\ &= 1 + \langle Ixz \rangle - \frac{1}{2} \langle (xz)^2 \rangle + \langle o(z^2) \rangle \\ &= 1 + I \langle x \rangle z - \frac{1}{2} \langle x^2 \rangle z^2 + \langle o(z^2) \rangle. \end{aligned} \quad (1.13)$$

Tomando o logaritmo, temos:

$$\ln(\psi_x(z)) = \ln\left(1 + I\langle x \rangle z - \frac{1}{2}mz^2 + \langle o(z^2) \rangle\right), \quad (1.14)$$

sendo $m = \langle x^2 \rangle$, mas, lembrando da expansão em série da função logaritmo, temos:

$$\begin{aligned} \ln(\psi_x(z)) &= I\langle x \rangle z - \frac{1}{2}mz^2 + \langle o(z^2) \rangle \\ &= I\langle x \rangle z - \frac{z^2}{2}m \left[1 - \frac{2\langle o(z^2) \rangle}{mz^2}\right]. \end{aligned} \quad (1.15)$$

Aplicando a exponencial:

$$\psi_x(z) = \exp\left\{I\langle x \rangle z - \frac{1}{2}mz^2 [1 + \omega_x(z)]\right\}, \quad \omega_x(z) = -2\frac{\langle o(z^2) \rangle}{mz^2}. \quad (1.16)$$

Note que, pela definição de $\omega_x(z)$, $\lim_{z \rightarrow 0} \omega_x(z) = \omega_x(0) = 0$. A forma da função característica em termos da função ω é chamada forma canônica de Lévy da função característica (LÉVY, 1924), que se apresentará ainda mais interessante quando considerarmos vetores aleatórios, por envolver a matriz de covariância das componentes do vetor, como veremos na próxima seção.

Podemos definir, em termos da variável aleatória original, a chamada variável centralizada, $x^{(c)} = x - \langle x \rangle$, de modo que $\langle x^{(c)} \rangle = 0$ e, portanto,

$$\psi_{x^{(c)}}(z) = \exp\left\{-\frac{z^2}{2}m^{(c)} [1 + \omega_{x^{(c)}}(z)]\right\}, \quad m^{(c)} = \langle x^2 \rangle - \langle x \rangle^2. \quad (1.17)$$

Podemos, ainda, após centralizar a variável aleatória, normalizá-la pelo seu desvio padrão, de modo que sua variância se torne unitária. Fazemos isso definindo a chamada variável reduzida:

$$x^{(r)} = \frac{x^{(c)}}{\sqrt{m^{(c)}}} = \frac{x - \langle x \rangle}{\sqrt{\langle x^2 \rangle - \langle x \rangle^2}}. \quad (1.18)$$

Agora, calculando a função característica da variável aleatória reduzida, temos:

$$\psi_{x^{(r)}}(z) = \exp\left\{-\frac{z^2}{2} [1 + \omega_{x^{(r)}}(z)]\right\}. \quad (1.19)$$

Esta é uma forma muito simples para a função característica, e que virá a facilitar as análises a serem feitas ao longo deste trabalho, especialmente por evidenciar a distribuição normal como seu caso especial, quando $\omega_{x(r)}(z) = 0 \quad \forall z$. Isto será fundamental para determinar a gaussianização da distribuição limite de um processo estocástico discreto.

Note que podemos escrever a função ω em termos da função característica ψ . Aqui, podemos escrever $\omega(z) = \omega_R(z) + I\omega_I(z)$ e $\psi(z) = \psi_R(z) + I\psi_I(z)$. Deste modo, temos:

$$\psi_R(z) + I\psi_I(z) = \exp \left[-\frac{z^2}{2}(1 + \omega_R(z) + I\omega_I(z)) \right]. \quad (1.20)$$

Tomando o módulo de ambos os lados:

$$|\psi(z)| = \exp \left[-\frac{z^2}{2}(1 + \omega_R(z)) \right] \Rightarrow 1 + \omega_R = -\frac{2}{z^2} \ln |\psi(z)|$$

$$\therefore \omega_R(z) = \frac{-2 \ln |\psi(z)| - z^2}{z^2}. \quad (1.21)$$

Note que z^2 e $\ln |\psi(z)|$ são pares e, portanto $\omega_R(z)$ é par. Agora, se tomamos a Equação (1.20), e dividimos os dois lados por seus módulos, obtemos:

$$\frac{\psi_R(z) + I\psi_I(z)}{|\psi(z)|} = \exp \left[-I\frac{z^2}{2}\omega_I(z) \right], \quad (1.22)$$

onde $z^2\omega_I(z)/2$ é o argumento da expressão do lado esquerdo. A solução é:

$$\omega_I(z) = -\frac{2}{z^2} \arctan2 \left(\frac{\psi_I(z)}{|\psi(z)|}, \frac{\psi_R(z)}{|\psi(z)|} \right). \quad (1.23)$$

A função $\arctan2(y, x)$ computa o valor principal do argumento do número complexo $x + Iy$, ou seja, $\arctan2(y, x) = \text{Arg}(x + Iy)$, no intervalo $(-\pi, \pi]$. Sua definição é:

$$\arctan2(y, x) = \begin{cases} \arctan(y/x), & \text{se } x > 0, \\ \arctan(y/x) + \pi, & \text{se } x < 0 \text{ e } y \geq 0, \\ \arctan(y/x) - \pi, & \text{se } x < 0 \text{ e } y < 0, \\ \pi/2, & \text{se } x = 0 \text{ e } y > 0, \\ -\pi/2, & \text{se } x = 0 \text{ e } y < 0, \\ \text{não definido}, & \text{se } x = 0 \text{ e } y = 0. \end{cases} \quad (1.24)$$

1.2 Vetores Aleatórios

Tendo conhecimento do que é uma variável aleatória, podemos definir um vetor aleatório n -dimensional $\mathbf{x} = (x_1, x_2, \dots, x_n)$, onde cada uma de suas componentes se comporta como uma variável aleatória, não necessariamente com as mesmas distribuições de probabilidades. Para um vetor aleatório, podemos definir sua função distribuição de probabilidades conjunta, que descreve todas as componentes do vetor, de modo que, para preservar sua normalização:

$$\int_{-\infty}^{\infty} d\mathbf{x} \rho_{\mathbf{x}}(\mathbf{x}) = \int_{-\infty}^{\infty} dx_1 \int_{-\infty}^{\infty} dx_2 \dots \int_{-\infty}^{\infty} dx_n \rho_{\mathbf{x}}(x_1, x_2, \dots, x_n) = 1. \quad (1.25)$$

Da mesma maneira, calculamos as probabilidades:

$$P(a_i \leq x_i \leq b_i \forall i \in \{1, 2, \dots, n\}) = \int_{a_1}^{b_1} dx_1 \int_{a_2}^{b_2} dx_2 \dots \int_{a_n}^{b_n} dx_n \rho_{\mathbf{x}}(\mathbf{x}). \quad (1.26)$$

Podemos escrever relações entre as probabilidades de diferentes componentes, através das chamadas probabilidades condicionais. $P(A|B)$ denota a probabilidade de um evento A dada a ocorrência de um evento B . Note que, então, a probabilidade de acontecerem A e B deve ser $P(A, B) = P(A|B)P(B) = P(B|A)P(A)$. A partir disto, podemos calcular a função distribuição de probabilidade de uma componente qualquer x_j do vetor \mathbf{x} , através do cálculo de suas probabilidades. Note:

$$P(a \leq x_j \leq b) = P(a \leq x_j \leq b, x_i \in \mathcal{R} \forall i \neq j) \Rightarrow \int_a^b dx_j \rho_j(x_j) = \int_{-\infty}^{\infty} dx_1 \dots \int_{-\infty}^{\infty} dx_{j-1} \int_a^b dx_j \int_{-\infty}^{\infty} dx_{j+1} \dots \int_{-\infty}^{\infty} dx_n \rho_{\mathbf{x}}(\mathbf{x}). \quad (1.27)$$

$$\therefore \rho_j(x_j) = \int_{-\infty}^{\infty} dx_1 \dots \int_{-\infty}^{\infty} dx_{j-1} \int_{-\infty}^{\infty} dx_{j+1} \dots \int_{-\infty}^{\infty} dx_n \rho_{\mathbf{x}}(x_1, \dots, x_j, \dots, x_n). \quad (1.28)$$

Ou seja, a função distribuição de probabilidade de uma variável específica dentro de um conjunto, chamada distribuição marginal, deve ser dada pela integração da função distribuição de probabilidades conjunta, com relação às outras variáveis que não aquela cuja distribuição é desejada. Este resultado decorre do fato apresentado na primeira linha: quando falamos da probabilidade de apenas uma variável dentre várias, considerar a probabilidade apenas da variável em questão, ignorando as outras, é o mesmo que dizer que as outras variáveis podem assumir qualquer um de seus possíveis valores.

Podemos calcular momentos estatísticos também para vetores aleatórios, é claro, pela definição dos valores médios de potências das componentes do vetor:

$$\langle x_j^m \rangle = \int_{-\infty}^{\infty} d\mathbf{x} x_j^m \rho_{\mathbf{x}}(\mathbf{x}). \quad (1.29)$$

Podemos calcular também momentos estatísticos mistos

$$\langle x_i^m x_j^l \rangle = \int_{-\infty}^{\infty} d\mathbf{x} x_i^m x_j^l \rho_{\mathbf{x}}(\mathbf{x}), \quad (1.30)$$

que são importantes para avaliar possíveis relações entre diferentes componentes do vetor aleatório. O caso mais utilizado é a covariância, dada por:

$$\text{cov}(x_i, x_j) = \langle (x_i - \langle x_i \rangle)(x_j - \langle x_j \rangle) \rangle = \langle x_i x_j \rangle - \langle x_i \rangle \langle x_j \rangle. \quad (1.31)$$

Se $i = j$, temos a variância, o segundo momento estatístico em torno da média.

1.2.1 A Forma Canônica da Função Característica de um Vetor Aleatório com Componentes de Variância Finita

Consideremos um vetor aleatório $\mathbf{x} = (x_1, x_2, \dots, x_n)$. Sua função característica é dada por:

$$\psi_{\mathbf{x}}(\mathbf{z}) = \langle \exp(I\mathbf{x} \cdot \mathbf{z}) \rangle = \left\langle \exp\left(I \sum_{i=1}^n x_i z_i\right) \right\rangle. \quad (1.32)$$

Vamos realizar o mesmo procedimento feito para o caso unidimensional. Escrevendo a série de Taylor da exponencial:

$$\begin{aligned} \psi_{\mathbf{x}}(\mathbf{z}) &= \left\langle 1 + I\mathbf{x} \cdot \mathbf{z} - \frac{1}{2} (\mathbf{x} \cdot \mathbf{z})^2 + o(z^2) \right\rangle \\ &= 1 + I \langle \mathbf{x} \rangle \cdot \mathbf{z} - \frac{1}{2} \sum_{i,j} \langle x_i x_j \rangle z_i z_j + \langle o(z^2) \rangle. \end{aligned} \quad (1.33)$$

Aqui, $\langle \mathbf{x} \rangle = (\langle x_1 \rangle, \langle x_2 \rangle, \dots, \langle x_n \rangle)$, ou seja, continua sendo um vetor, com suas componentes sendo as médias de cada componente do vetor aleatório original. Note:

$$\sum_{i,j}^n \langle x_i x_j \rangle z_i z_j = \sum_{i,j}^n z_i \langle x_i x_j \rangle z_j = \mathbf{z} \mathfrak{m} \mathbf{z}, \quad \mathfrak{m}_{i,j} = \langle x_i x_j \rangle, \quad (1.34)$$

ou seja, \mathfrak{m} é a matriz dos momentos estatísticos de segunda ordem do vetor aleatório \mathbf{x} . Continuando, tomamos o logaritmo:

$$\ln(\psi_{\mathbf{x}}(\mathbf{z})) = \ln \left(1 + I \langle \mathbf{x} \rangle \cdot \mathbf{z} - \frac{1}{2} \mathbf{z} \mathfrak{m} \mathbf{z} + \langle o(z^2) \rangle \right). \quad (1.35)$$

Expandindo em série de potências novamente:

$$\begin{aligned} \ln(\psi_{\mathbf{x}}(\mathbf{z})) &= I \langle \mathbf{x} \rangle \cdot \mathbf{z} - \frac{1}{2} \mathbf{z} \mathfrak{m} \mathbf{z} + \langle o(z^2) \rangle \\ &= I \langle \mathbf{x} \rangle \cdot \mathbf{z} - \frac{1}{2} \mathbf{z} \mathfrak{m} \mathbf{z} [1 + \omega_{\mathbf{x}}(\mathbf{z})], \quad \omega_{\mathbf{x}}(\mathbf{z}) = -2 \frac{\langle o(z^2) \rangle}{\mathbf{z} \mathfrak{m} \mathbf{z}}. \end{aligned} \quad (1.36)$$

Novamente definimos a função ω , desta vez associada a um vetor. Dada sua definição, espera-se novamente que $\omega_{\mathbf{x}}(\mathbf{z}) \rightarrow \omega_{\mathbf{x}}(\mathbf{0}) = 0$, para $\mathbf{z} \rightarrow 0$. Portanto, a função característica de \mathbf{x} é dada por:

$$\psi_{\mathbf{x}}(\mathbf{z}) = \exp \left\{ I \langle \mathbf{x} \rangle \cdot \mathbf{z} - \frac{1}{2} \mathbf{z} \mathfrak{m} \mathbf{z} [1 + \omega_{\mathbf{x}}(\mathbf{z})] \right\}. \quad (1.37)$$

Para o vetor centralizado $\mathbf{x}^{(c)} = \mathbf{x} - \langle \mathbf{x} \rangle$, temos:

$$\psi_{\mathbf{x}^{(c)}}(\mathbf{z}) = \exp \left\{ -\frac{1}{2} \mathbf{z} \mathfrak{m}^{(c)} \mathbf{z} [1 + \omega_{\mathbf{x}^{(c)}}(\mathbf{z})] \right\}, \quad \mathfrak{m}_{i,j}^{(c)} = \langle x_i x_j \rangle - \langle x_i \rangle \langle x_j \rangle. \quad (1.38)$$

Note que $\mathfrak{m}^{(c)}$ é a matriz de covariância do vetor \mathbf{x} . Definindo em seguida o vetor centralizado e normalizado $\mathbf{x}^{(r)}$, ou seja, com suas componentes normalizadas, de modo que a variância de cada uma seja unitária, temos:

$$x_i^{(r)} = \frac{x_i^{(c)}}{\sqrt{m_{i,i}^{(c)}}} = \frac{x_i - \langle x_i \rangle}{\sqrt{\langle x_i^2 \rangle - \langle x_i \rangle^2}}. \quad (1.39)$$

Com isto, a função característica é dada por

$$\psi_{\mathbf{x}^{(r)}}(\mathbf{z}) = \exp \left\{ -\frac{1}{2} \mathbf{z} \mathfrak{m}^{(r)} \mathbf{z} [1 + \omega_{\mathbf{x}^{(r)}}(\mathbf{z})] \right\}, \quad (1.40)$$

sendo $m^{(r)}$ a matriz de correlação de \mathbf{x} :

$$m_{i,j}^{(r)} = \text{corr}(x_i, x_j) = \frac{\langle x_i x_j \rangle - \langle x_i \rangle \langle x_j \rangle}{\sqrt{(\langle x_i^2 \rangle - \langle x_i \rangle^2)(\langle x_j^2 \rangle - \langle x_j \rangle^2)}} \quad (1.41)$$

1.2.2 O Vetor Principal de um Vetor Aleatório

Por definição, a matriz m – assim como $m^{(c)}$ e $m^{(r)}$ – é simétrica. O que significa que ela pode ser diagonalizada e todos os seus autovalores são reais. A transformação para a forma diagonal é dada por uma matriz ortogonal \mathbb{T} , o que significa $\mathbb{T}\bar{\mathbb{T}} = \bar{\mathbb{T}}\mathbb{T} = \mathbb{1}$, ou seja, $\mathbb{T}^{-1} = \bar{\mathbb{T}}$, sendo $\bar{\mathbb{T}}$ a matriz transposta de \mathbb{T} . Chamamos, então, de vetor principal, o vetor aleatório após sofrer a transformação que diagonaliza sua matriz de covariância.

Consideremos um vetor centralizado $\mathbf{x}^{(c)}$ com matriz de covariância $m^{(c)}$, tal que $\mathbb{T}m\bar{\mathbb{T}} = \mathbb{D}_{m^{(c)}} = \text{diag}(m_1^{(c)}, m_2^{(c)}, \dots, m_n^{(c)})$. Como a matriz de covariância deve ser positivo definida, todos os seus autovalores são positivos. Definindo o vetor principal:

$$\mathbf{x}^{(p)} = \mathbb{T}\mathbf{x}^{(c)}, \quad (1.42)$$

sua função característica deve ser dada por:

$$\psi_{\mathbf{x}^{(p)}}(\mathbf{z}) = \exp \{ I_{\mathbf{x}^{(p)}} \cdot \mathbf{z} \} = \exp \{ I(\mathbb{T}\mathbf{x}^{(c)}) \cdot \mathbf{z} \} = \exp \{ I_{\mathbf{x}^{(c)}} \cdot (\bar{\mathbb{T}}\mathbf{z}) \} \quad (1.43)$$

$$\therefore \psi_{\mathbf{x}^{(p)}}(\mathbf{z}) = \psi_{\mathbf{x}^{(c)}}(\bar{\mathbb{T}}\mathbf{z}). \quad (1.44)$$

Lembrando da função característica de $\mathbf{x}^{(c)}$ (1.38), podemos escrever para o vetor principal:

$$\begin{aligned} \psi_{\mathbf{x}^{(p)}}(\mathbf{z}) &= \exp \left\{ -\frac{1}{2} (\bar{\mathbb{T}}\mathbf{z}) m^{(c)} (\bar{\mathbb{T}}\mathbf{z}) [1 + \omega_{\mathbf{x}^{(c)}}(\bar{\mathbb{T}}\mathbf{z})] \right\} \\ &= \exp \left\{ -\frac{1}{2} \mathbf{z} \mathbb{T} m^{(c)} \bar{\mathbb{T}} \mathbf{z} [1 + \omega_{\mathbf{x}^{(c)}}(\bar{\mathbb{T}}\mathbf{z})] \right\} \\ &= \exp \left\{ -\frac{1}{2} \mathbf{z} \mathbb{D}_{m^{(c)}} \mathbf{z} [1 + \omega_{\mathbf{x}^{(p)}}(\mathbf{z})] \right\}, \quad \omega_{\mathbf{x}^{(p)}}(\mathbf{z}) = \omega_{\mathbf{x}^{(c)}}(\bar{\mathbb{T}}\mathbf{z}). \end{aligned} \quad (1.45)$$

Podemos passar agora ao vetor principal centralizado e normalizado, $\mathbf{x}^{(r)}$, utilizando as raízes dos autovalores de $m^{(c)}$ como normalizações:

$$x_i^{(r)} = \frac{x_i^{(p)}}{\sqrt{m_i^{(c)}}} \quad (1.46)$$

Com isto, lembrando que podemos reescrever:

$$\mathbf{z} \mathbb{D}_{\mathbf{m}^{(c)}} \mathbf{z} = \sum_{i,j=1}^n (\mathbb{D}_{\mathbf{m}^{(c)}})_{i,j} z_i z_j = \sum_{i,j=1}^n m_i^{(c)} \delta_{i,j} z_i z_j = \sum_{i=1}^n m_i^{(c)} z_i^2, \quad (1.47)$$

calculamos a função característica:

$$\begin{aligned} \psi_{\mathbf{x}^{(r)}}(\mathbf{z}) &= \exp \left\{ I \sum_{i=1}^n x_i^{(r)} z_i \right\} = \exp \left\{ I \sum_{i=1}^n \frac{x_i^{(p)}}{\sqrt{m_i^{(c)}}} z_i \right\} = \exp \left\{ I \sum_{i=1}^n x_i^{(p)} \frac{z_i}{\sqrt{m_i^{(c)}}} \right\} \\ &= \exp \left\{ -\frac{1}{2} \left[\sum_{i,j=1}^n (\mathbb{D}_{\mathbf{m}^{(c)}})_{i,j} \frac{z_i}{\sqrt{m_i^{(c)}}} \frac{z_j}{\sqrt{m_j^{(c)}}} \right] [1 + \omega_{\mathbf{x}^{(r)}}(\mathbf{z})] \right\} \\ &= \exp \left\{ -\frac{1}{2} \left[\sum_{i=1}^n m_i^{(c)} \frac{z_i^2}{m_i^{(c)}} \right] [1 + \omega_{\mathbf{x}^{(r)}}(\mathbf{z})] \right\} \\ &= \exp \left\{ -\frac{1}{2} \mathbf{z} \cdot \mathbf{z} [1 + \omega_{\mathbf{x}^{(r)}}(\mathbf{z})] \right\}, \text{ sendo} \end{aligned} \quad (1.48)$$

$$\omega_{\mathbf{x}^{(r)}}(\mathbf{z}) = \omega_{\mathbf{x}^{(p)}} \left(\left\{ \frac{z_i}{\sqrt{m_i^{(c)}}}, i = 1, 2, \dots, n \right\} \right) \quad (1.49)$$

Concluimos que a matriz de covariância do vetor principal $\mathbf{x}^{(r)}$ é a matriz identidade.

1.2.3 Vetor Principal de Componentes Independentes

Agora, vamos verificar o caso especial em que o vetor principal possui componentes independentes entre si. Aqui, vamos observar que podemos separar a função ω de cada componente. Escrevendo a função característica, a independência das componentes nos permite escrever o produtório das funções características de cada componente:

$$\begin{aligned} \psi_{\mathbf{x}^{(p)}}(\mathbf{z}) &= \langle e^{I \mathbf{x}^{(p)} \cdot \mathbf{z}} \rangle = \left\langle \exp \left\{ I \sum_{i=1}^n x_i^{(p)} z_i \right\} \right\rangle = \prod_{i=1}^n \langle \exp (I x_i^{(p)} z_i) \rangle \\ &= \prod_{i=1}^n \psi_{x_i^{(p)}}(z_i). \end{aligned} \quad (1.50)$$

O mesmo vale para o vetor principal normalizado, é claro:

$$\psi_{\mathbf{x}^{(r)}}(\mathbf{z}) = \prod_{i=1}^n \psi_{x_i^{(r)}}(z_i) = \prod_{i=1}^n \exp \left\{ -\frac{1}{2} z_i^2 [1 + \omega_i(z_i)] \right\}, \quad \omega_i = \omega_{x_i^{(r)}}. \quad (1.51)$$

Deste modo,

$$\begin{aligned} \psi_{\mathbf{x}^{(r)}}(\mathbf{z}) &= \exp \left\{ -\frac{1}{2} \sum_{i=1}^n z_i^2 [1 + \omega_i(z_i)] \right\} \\ &= \exp \left\{ -\frac{\mathbf{z} \cdot \mathbf{z}}{2} [1 + \omega_{\mathbf{x}^{(r)}}(\mathbf{z})] \right\}, \quad \omega_{\mathbf{x}^{(r)}}(\mathbf{z}) = \frac{1}{\mathbf{z} \cdot \mathbf{z}} \sum_{i=1}^n z_i^2 \omega_i(z_i). \end{aligned} \quad (1.52)$$

Como esperado, caso as componentes do vetor principal sejam independentes, conseguimos escrever a função ω como a soma das funções de cada uma das componentes. Vale notar que, se as componentes forem identicamente distribuídas, temos $\omega_i = \omega \quad \forall i$.

1.3 A Função Característica de Paul Lévy Para Somas de Variáveis Aleatórias

Consideremos a soma X_n de variáveis aleatórias x_i :

$$X_n = \sum_{i=1}^n x_i, \quad (1.53)$$

onde cada variável aleatória x_i tem sua distribuição de probabilidades dada por $f_i(x_i)$. Esta exposição sobre a função característica a seguir é inspirada em (FIGUEIREDO et al., 2006), que serviu de base para o estudo do método, mas vai além do que é apresentado aqui.

Podemos calcular os momentos estatísticos iniciais para as variáveis aleatórias x_i , assim como para sua soma X_n :

$$\begin{cases} a_i = \langle x_i \rangle; \quad A_n = \langle X_n \rangle; \\ m_i^2 = \langle x_i^2 \rangle - a_i^2; \quad M_n^2 = \langle X_n^2 \rangle - A_n^2. \end{cases} \quad (1.54)$$

Vamos avaliar o caso em que as variáveis aleatórias x_i são independentes entre si, de modo que sua covariância seja nula:

$$\langle x_i x_j \rangle = a_i a_j, \quad \forall i \neq j. \quad (1.55)$$

Desta maneira, temos uma forma simples para a variância da soma X_n :

$$\begin{aligned} M_n^2 &= -A_n^2 + \left\langle \left[\sum_{i=1}^n x_i \right]^2 \right\rangle = -A_n^2 + \sum_{i=1}^n \langle x_i^2 \rangle + 2 \sum_{i < j} \langle x_i x_j \rangle \\ &= -A_n^2 + \sum_{i=1}^n (m_i^2 + a_i^2) + 2 \sum_{i < j} a_i a_j \end{aligned} \quad (1.56)$$

Note que

$$A_n^2 = \langle X_n \rangle^2 = \left(\sum_{i=1}^n a_i \right)^2 = \sum_{i=1}^n a_i^2 + 2 \sum_{i < j} a_i a_j. \quad (1.57)$$

$$\therefore M_n^2 = \sum_{i=1}^n m_i^2 \quad (1.58)$$

Definimos, então, as variáveis reduzidas associadas à nossa soma:

$$x_i^{(r)} = \frac{x_i - a_i}{m_i}; \quad (1.59)$$

$$X_n^{(r)} = \frac{X_n - A_n}{M_n} = \frac{-A_n}{M_n} + \frac{1}{M_n} \sum_{i=1}^n (m_i x_i^{(r)} + a_i). \quad (1.60)$$

Porém, note que:

$$A_n = \langle X_n \rangle = \sum_{i=1}^n \langle x_i \rangle = \sum_{i=1}^n a_i, \quad (1.61)$$

e, portanto:

$$X_n^{(r)} = \sum_{i=1}^n \frac{m_i}{M_n} x_i^{(r)}. \quad (1.62)$$

As funções características das variáveis reduzidas $x_i^{(r)}$ estão diretamente associadas às aquelas de suas variáveis normais, e uma pode ser escrita em termos da outra:

$$\psi_i^{(r)}(z) = \left\langle \exp \left[Iz \frac{x_i - a_i}{m_i} \right] \right\rangle = \psi_i \left(\frac{z}{m_i} \right) \exp \left[-Iz \frac{a_i}{m_i} \right]; \quad (1.63)$$

$$\psi_i(z) = \psi_i^{(r)}(m_i z) \exp [I z a_i]. \quad (1.64)$$

Podemos, agora, calcular a função característica da variável reduzida $X_n^{(r)}$:

$$\Psi_n^{(r)}(z) = \left\langle \exp \left[I \frac{z}{M_n} \sum_{i=1}^n m_i x_i^{(r)} \right] \right\rangle = \left\langle \prod_{i=1}^n \exp \left[I z \frac{m_i x_i^{(r)}}{M_n} \right] \right\rangle.$$

Considerando que as variáveis $x_i^{(r)}$ são independentes entre si, e portanto $\langle (x_i^{(r)})^n (x_j^{(r)})^m \rangle = \delta_{i,j} \mu_{i,n+m} / \mu_{i,2}$, sendo $\mu_{i,n}$ o n -ésimo momento estatístico de x_i , temos:

$$\begin{aligned} \Psi_n^{(r)}(z) &= \left\langle \prod_{i=1}^n \left[\sum_{j=1}^{\infty} \frac{1}{j!} \left(I z \frac{m_i x_i^{(r)}}{M_n} \right)^j \right] \right\rangle = \prod_{i=1}^n \left\langle \exp \left[I z \frac{m_i x_i^{(r)}}{M_n} \right] \right\rangle \\ &\therefore \Psi_n^{(r)}(z) = \prod_{i=1}^n \psi_i^{(r)} \left(\frac{m_i}{M_n} z \right), \end{aligned} \quad (1.65)$$

ou seja, podemos escrever a função característica de $X_n^{(r)}$, $\Psi_n^{(r)}(z)$, em termos das funções características das variáveis $x_i^{(r)}$, $\psi_i^{(r)}(z)$. Baseado nisto, apresentamos a seguir um teorema sobre a distribuição limite de uma soma de variáveis aleatórias, teorema do qual o Teorema do Limite Central é apenas um caso especial.

1.3.1 Generalizando o Teorema do Limite Central de Paul Lévy

Apresentamos, a seguir, um teorema mais geral sobre distribuições estáveis para somas de variáveis aleatórias, que tem o TLC como um caso especial.

Teorema 1.1. *Considere um conjunto de variáveis aleatórias x_i , $i \in 1, 2, \dots, n$, sendo m_i o desvio padrão de x_i e M_n o desvio padrão da soma $X_n = x_1 + x_2 + \dots + x_n$. Se:*

1. $\exists \lim_{n \rightarrow \infty} (m_i / M_n) \forall i$;
2. $\exists L_z, |\omega_i(z')| \leq L_z, \forall i, \forall z', -z < z' < z$;
3. $\omega_i(z)$ é contínua em $z = 0$;

então:

$$\exists \lim_{n \rightarrow \infty} \Psi_n^{(r)}(z) = \Psi^{(r)}(z) = \exp \left[-\frac{z^2}{2} (1 + \Omega(z)) \right], \quad \Omega(z) = \sum_{i=1}^{\infty} \left[\frac{m_i^2}{M_n^2} \omega_i \left(\frac{m_i}{M_n} z \right) \right],$$

onde $\Psi^{(r)}(z)$ é a função característica da variável reduzida da soma $X_n^{(r)}$.

Note que, de acordo com o Teorema 1.1, a função característica (e, portanto, a função distribuição de probabilidades) da soma de variáveis só será gaussiana caso as razões $(m_i/M_n)^2 \rightarrow 0$ quando $n \rightarrow \infty$ – ou no caso em que $\omega_i(z) = 0 \forall z$, o que acontece apenas se todas as variáveis somadas têm distribuições gaussianas. Caso as variáveis não sejam gaussianas, é claro que a função característica dependerá das funções $\omega_i(z)$ e não será gaussiana. No Capítulo 2, veremos um exemplo muito simples para o qual a função característica não converge para uma gaussiana.

A demonstração do Teorema 1.1 pode ser encontrada no Apêndice A. Detalhamos, a seguir, o cálculo a função $\Omega(z)$ em termos das funções $\omega_i(z)$, associadas às funções características de cada variável aleatória da soma.

Lembrando da Equações (1.65) e (1.18), podemos escrever:

$$\begin{aligned} \Psi_n^{(r)}(z) &= \prod_{i=1}^n \exp \left[-\frac{1}{2} \frac{m_i^2}{M_n^2} z^2 \left(1 + \omega_i \left(\frac{m_i}{M_n} z \right) \right) \right] = \exp \left\{ -\frac{z^2}{2} \sum_{i=1}^n \left[\frac{m_i^2}{M_n^2} \left(1 + \omega_i \left(\frac{m_i}{M_n} z \right) \right) \right] \right\} \\ \therefore \Psi_n^{(r)}(z) &= \exp \left[-\frac{z^2}{2} (1 + \Omega_n(z)) \right], \end{aligned} \quad (1.66)$$

onde $\Omega_n(z)$ é:

$$\Omega_n(z) = \sum_{i=1}^n \left[\frac{m_i^2}{M_n^2} \omega_i \left(\frac{m_i}{M_n} z \right) \right]. \quad (1.67)$$

Considerando a hipótese 1 do Teorema 1.1, podemos tomar o limite $n \rightarrow \infty$ de $\Omega_n(z)$:

$$\Omega(z) = \lim_{n \rightarrow \infty} \Omega_n(z) = \sum_{i=1}^{\infty} \lambda_i^2 \omega_i(\lambda_i z); \quad \lambda_i = \lim_{n \rightarrow \infty} \frac{m_i}{M_n}. \quad (1.68)$$

Note que se

$$\forall i, \lim_{n \rightarrow \infty} \lambda_i = 0, \quad (1.69)$$

então a função característica $\Psi^{(r)}(z)$ converge para uma gaussiana e, como a transformada de Fourier de uma gaussiana também é gaussiana, a função distribuição $F^{(r)}(X_n^{(r)})$ deve convergir para a distribuição normal, e assim temos o Teorema do Limite Central como um caso especial do Teorema 1.1. A expressão (1.69) é a hipótese da infinitesimalidade (das razões m_i/M_n), e a forma da função característica expressa nas Equações (1.66) e (1.67) explicita muito bem a relação de necessidade da hipótese de infinitesimalidade para validade do TLC.

1.4 Expansão em Série de Potências da Função Característica

Sabemos que podemos escrever a função característica de uma variável aleatória em termos de uma série de potências associada aos momentos estatísticos da variável aleatória em questão (1.12). Da mesma maneira, a função característica de uma soma de n variáveis aleatórias é dada em termo de seus momentos estatísticos (aqui dados por $\nu_{n,j}$):

$$\Psi_n(z) = \sum_{j=0}^{\infty} \frac{(Iz)^j}{j!} \nu_{n,j}. \quad (1.70)$$

Lembrando que a função característica da soma é dada pelo produtório das funções características das variáveis, temos:

$$\sum_{j=0}^{\infty} \frac{(Iz)^j}{j!} \nu_{n,j} = \prod_{i=1}^n \left[\sum_{j=1}^{\infty} \frac{(Iz)^j}{j!} \mu_{i,j} \right], \quad (1.71)$$

onde $\mu_{i,j}$ é o j -ésimo momento estatístico da i -ésima variável aleatória. Esta relação é conveniente para o cálculo de momentos estatísticos de baixa ordem da soma de variáveis. Se estamos considerando variáveis centralizadas, $\mu_{i,1} = 0$ (o que pode ser feito por uma simples transformação $x'_i = x_i - \mu_{i,1} \rightarrow X'_n = X_n - \mu_{1,1} - \mu_{2,1} - \dots - \mu_{n,1}$), obtemos:

$$\nu_{n,2} = \sum_{i=1}^n \mu_{i,2}; \quad \nu_{n,3} = \sum_{i=1}^n \mu_{i,3}.$$

Nos casos dos momentos de ordem mais alta, aparecerem termos mistos, envolvendo produtos de momentos estatísticos de variáveis diferentes, dadas as diferentes combinações na Equação (1.71) que podem gerar uma determinada potência de z . Por exemplo, para z^4 , temos:

$$\nu_{n,4} = \sum_{i=1}^n \mu_{i,4} + \frac{4!}{2!2!} \sum_{i=1}^{n-1} \sum_{j=i+1}^n \mu_{i,2} \cdot \mu_{j,2}. \quad (1.72)$$

Temos uma simplificação da expansão em séries da função característica da soma de variáveis aleatórias quando consideramos que todas as variáveis possuem a mesma distribuição, quando reduzidas, como será visto a seguir.

1.4.1 Variáveis Reduzidas Identicamente Distribuídas

Sabemos que, para qualquer variável aleatória, podemos definir sua variável reduzida (1.59), que possui, obrigatoriamente, média nula e desvio padrão unitário. No caso

em que todas as variáveis reduzidas $x_i^{(r)}$ possuem a mesma densidade de probabilidade, dizemos que temos variáveis reduzidas identicamente distribuídas.

Note que a identidade entre as distribuições das variáveis reduzidas implica na identidade de suas funções características e, portanto, $\omega_i(z) = \omega(z) \forall i$. Será mostrado a seguir como a consideração de um processo com variáveis reduzidas identicamente distribuídas leva a uma simplificação considerável do problema, facilitando o cálculo das funções $\omega(z)$ e $\Omega(z)$. As funções características de todas as variáveis reduzidas x_i serão dadas por:

$$\psi_i^{(r)}(z) = \psi^{(r)}(z) = \exp \left[-\frac{z^2}{2}(1 + \omega(z)) \right] \forall i. \quad (1.73)$$

Vamos considerar o caso em que todas as variáveis x_i são centralizadas, ou seja, $\mu_{i,1} = a_i = 0 \forall i$. Utilizando as Equações (1.63) e (1.12), se definimos $\nu_{i,j} = \mu_{i,j}/m_i^j$, sendo $\mu_{i,j} = \langle x_i^j \rangle$ e $m_i = \mu_{i,2}$, temos:

$$\psi_i^{(r)}(z) = \sum_{n=0}^{\infty} \frac{(Iz)^n}{n!} \nu_{i,n}, \quad (1.74)$$

porém se as funções características são iguais para todas as variáveis reduzidas, podemos dizer que $\nu_{i,n} = \nu_n \forall i$. Podemos escrever a função $\omega(z)$ também por uma séries de potências. Sabendo que sua parte real é par e sua parte imaginária, ímpar, temos:

$$\omega(z) = \omega_R(z) + I\omega_I(z) = \sum_{n=1}^{\infty} K_n \frac{(Iz)^n}{n!}. \quad (1.75)$$

Note que o termo ($n = 0$) não aparece, pois sabemos que $\omega(0) = 0$. Deste modo:

$$K_n = (-I)^n \left. \frac{d^n}{dz^n} \omega(z) \right|_{z=0} \quad (1.76)$$

Utilizando as Equações (1.74) e (1.75), é fácil determinar K_n em termos de ν_n . O processo é extremamente simples, porém igualmente desinteressante, uma vez que consiste apenas em derivar a função característica na formas (1.73) (substituindo $\omega(z)$ pela forma (1.75)) e (1.74), comparando os coeficientes de cada potência dos resultados.

Foi desenvolvido, em Python – utilizando a biblioteca *sympy*, que permite a programação de matemática simbólica –, um *script* para o cálculo dos coeficientes K_i em termos dos momentos estatísticos ν_i , disponibilizado num repositório online.¹

¹ O programa está disponível no GitHub: <https://github.com/joaopedro-vm/PID_omega_coeffs>.

Da mesma maneira que foi feito para $\omega(z)$, podemos escrever a expansão em série de potências para a função $\Omega(z)$. Esta também deve ter parte real par, e parte imaginária ímpar:

$$\Omega(z) = \Omega_R(z) + I\Omega_I(z) = \sum_{n=1}^{\infty} L_n \frac{(Iz)^n}{n!}. \quad (1.77)$$

Substituindo as Equações (1.75) e (1.77) em (1.68), temos:

$$\sum_{n=1}^{\infty} L_n \frac{(Iz)^n}{n!} = \sum_{i=1}^{\infty} \lambda_i^2 \left[\sum_{n=1}^{\infty} K_n \frac{(I\lambda_i z)^n}{n!} \right] \Rightarrow \sum_{n=0}^{\infty} \frac{(Iz)^n}{n!} \left[L_n - K_n \sum_{i=1}^{\infty} \lambda_i^{n+2} \right] = 0.$$

$$\therefore L_n = K_n \sum_{i=1}^{\infty} \lambda_i^{n+2}, \quad \lambda_i = \frac{m_i}{M_n}. \quad (1.78)$$

Portanto, no caso de variáveis centralizadas, cujas variáveis reduzidas são identicamente distribuídas, obtemos uma forma fechada para a função característica da soma das variáveis aleatórias. Porém, isto não quer dizer que seu cálculo seja simples, uma vez que o cálculo de K_n por si só pode se tornar bastante intenso, mesmo para um computador, e pode ser necessária uma grande quantidade de termos da expansão de Ω para a determinação da função característica com precisão.

1.5 Potências de Matrizes

Dado o problema (13), é óbvio que precisamos calcular potências da matriz \mathbb{A} . É muito simples calcular potências de matrizes diagonalizáveis, uma vez que estas são similares a matrizes diagonais. Dizer que duas matrizes \mathbb{A} e \mathbb{B} são similares é dizer que existe uma matriz \mathbb{P} invertível (ou seja, existe matriz \mathbb{P}^{-1} , $\mathbb{P}\mathbb{P}^{-1} = \mathbb{P}^{-1}\mathbb{P} = \mathbb{1}$) tal que:

$$\mathbb{A} = \mathbb{P}^{-1}\mathbb{B}\mathbb{P}, \text{ equivalente a } \mathbb{P}\mathbb{A} = \mathbb{B}\mathbb{P}; \quad \mathbb{B} = \mathbb{P}\mathbb{A}\mathbb{P}^{-1}; \quad \mathbb{A}\mathbb{P}^{-1} = \mathbb{P}^{-1}\mathbb{B}. \quad (1.79)$$

Em nosso caso, que desejamos diagonalizar uma matriz \mathbb{M} de formato $n \times n$. Assim é possível mostrar que a solução é:

$$\mathbb{M} = \mathbb{P}\mathbb{D}\mathbb{P}^{-1}, \quad (1.80)$$

onde $\mathbb{D} = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_n)$, sendo λ_i o i -ésimo autovalor da matriz \mathbb{M} , e a matriz \mathbb{P} tem como sua i -ésima coluna o autovetor associado ao autovalor λ_i .

A diagonalização de uma matriz é muito útil para o cálculo de suas potências, pois dada a álgebra de multiplicação de matrizes, não é óbvio qual deve ser a n -ésima potência de uma matriz genérica, mas para uma matriz diagonal, a operação se torna muito simples, e se $\mathbb{D} = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_n)$, então $\mathbb{D}^m = \text{diag}(\lambda_1^m, \lambda_2^m, \dots, \lambda_n^m)$, para um m natural. Sabendo disso, calculamos a n -ésima potência de \mathbb{M} :

$$\mathbb{M}^n = \mathbb{P} \underbrace{\overbrace{\mathbb{D} \mathbb{P}^{-1} \mathbb{P} \mathbb{D} \mathbb{P}^{-1} \mathbb{P} \dots \mathbb{D} \mathbb{P}^{-1} \mathbb{P} \mathbb{D} \mathbb{P}^{-1}}^{= 1}}_{(n-1) \text{ vezes}} \mathbb{P} = \mathbb{P} \mathbb{D}^n \mathbb{P}^{-1}. \quad (1.81)$$

No caso em que a matriz \mathbb{M} não é diagonalizável, podemos ainda simplificar qualquer matriz substancialmente, facilitando o cálculo de suas potências, como será apresentado a seguir.

1.5.1 A Forma Canônica de Jordan Para Matrizes Não Diagonalizáveis

De acordo com (BRONSON, 1991), é possível demonstrar que qualquer matriz é similar a uma forma quase diagonal, esta forma dada por:

$$\mathbb{J} = \begin{bmatrix} \mathbb{J}_1 & & & & \\ & \mathbb{J}_2 & & \mathbf{0} & \\ & & \ddots & & \\ & \mathbf{0} & & \mathbb{J}_{n-1} & \\ & & & & \mathbb{J}_n \end{bmatrix}, \quad (1.82)$$

sendo cada \mathbb{J}_i , na verdade, também uma matriz:

$$\mathbb{J}_i = \underbrace{\begin{bmatrix} \lambda_i & 1 & 0 & \dots & 0 \\ 0 & \lambda_i & 1 & \ddots & \vdots \\ 0 & 0 & \ddots & \ddots & 0 \\ \vdots & \ddots & \ddots & \lambda_i & 1 \\ 0 & \dots & 0 & 0 & \lambda_i \end{bmatrix}}_{N_i \text{ termos}}, \quad (1.83)$$

onde λ_i é autovalor de \mathbb{J} com multiplicidade N_i . Portanto, se \mathbb{J} é uma matriz $L \times L$, então $N_1 + N_2 + \dots + N_n = L$. À forma (1.82), é dado o nome de forma de Jordan, e suas submatrizes \mathbb{J}_i são denominadas blocos de Jordan. É interessante notar que o formato (1.82) implica numa simplificação de suas potências:

$$\mathbb{J}^m = \begin{bmatrix} \mathbb{J}_1^m & & & \\ & \mathbb{J}_2^m & & \\ & & \ddots & \\ & & & \mathbb{J}_{n-1}^m \\ & 0 & & & \mathbb{J}_n^m \end{bmatrix}, \quad (1.84)$$

Mas nos resta calcular a forma geral de uma potência de um bloco de Jordan (1.83). Aqui surge uma leve complicação, cuja solução vamos apresentar a seguir. Note que podemos separar um bloco de Jordan $L \times L$ em duas matrizes:

$$\mathbb{J} = \begin{bmatrix} \lambda & 1 & \cdots & 0 \\ 0 & \lambda & \ddots & \vdots \\ \vdots & \ddots & \ddots & 1 \\ 0 & \cdots & 0 & \lambda \end{bmatrix} = \begin{bmatrix} \lambda & 0 & \cdots & 0 \\ 0 & \lambda & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \lambda \end{bmatrix} + \begin{bmatrix} 0 & 1 & \cdots & 0 \\ 0 & 0 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 1 \\ 0 & \cdots & 0 & 0 \end{bmatrix} = \lambda \mathbb{1}_L + \mathbb{N}_L, \quad (1.85)$$

onde $\mathbb{1}_L$ é a matriz identidade $L \times L$ e \mathbb{N}_L é o que podemos chamar de matriz superdiagonal, também $L \times L$. É importante notar que \mathbb{N}_L é uma matriz nilpotente, fato responsável por permitir que obtenhamos uma fórmula geral para a potência de um bloco de Jordan. Note que:

$$\mathbb{N}_L^{i,j} = \delta_{i,j-1}, \quad i \in \{1, 2, \dots, L-1\}. \quad (1.86)$$

$$\therefore (\mathbb{N}_L^2)^{i,j} = \sum_{k=1}^L \mathbb{N}_L^{i,k} \mathbb{N}_L^{k,j} = \sum_{k=1}^L \delta_{i,k-1} \delta_{k,l-1} = \delta_{i,l-2}, \quad (1.87)$$

ou seja, se \mathbb{N}_L é nula, com 1ª faixa acima da diagonal preenchida por 1's, então \mathbb{N}_L^2 tem a 2ª faixa acima da diagonal preenchida por 1's. De modo geral, então, obtemos:

$$(\mathbb{N}_L^n)^{i,j} = \delta_{i,j-n}, \quad (1.88)$$

com a n -ésima faixa acima da diagonal preenchida por 1's. Portanto, $\mathbb{N}_L^n = 0 \forall n \geq L$.

Outro fato importante para que o cálculo de potências de blocos de Jordan tenham fórmula fechada é que a matriz identidade $\mathbb{1}$ comuta com qualquer outra matriz, de modo que as potências $(\lambda \mathbb{1}_L + \mathbb{N}_L)^n$ pode ser calculada através do Teorema Binomial, de acordo com o qual, temos:

$$(\lambda \mathbb{1}_L + \mathbb{N}_L)^n = \sum_{k=0}^n \binom{n}{k} \lambda^{n-k} \mathbb{1}_L \mathbb{N}_L^k = \sum_{k=0}^W \binom{n}{k} \lambda^{n-k} \mathbb{N}_L^k, \quad W = \min\{L-1, n\} \quad (1.89)$$

Lembrando que $\mathbb{N}_L^0 = \mathbb{1}_L$, temos, de modo geral:

$$\mathbb{J}_L^n = \begin{bmatrix} \lambda^n & \binom{n}{1} \lambda^{n-1} & \binom{n}{2} \lambda^{n-2} & \dots & \dots & \binom{n}{L-1} \lambda^{n-L+1} \\ 0 & \lambda^n & \binom{n}{1} \lambda^{n-1} & \dots & \dots & \binom{n}{L-2} \lambda^{n-L+2} \\ 0 & 0 & \lambda^n & \ddots & \vdots & \vdots \\ 0 & 0 & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 & \lambda^n & \binom{n}{1} \lambda^{n-1} \\ 0 & \dots & 0 & 0 & 0 & \lambda^n \end{bmatrix}_{L \times L}. \quad (1.90)$$

Agora, basta fazer isso para cada bloco de Jordan da matriz desejada, e a potência de uma matriz na forma de Jordan será uma diagonal com cada entrada sendo a potência de um bloco de Jordan. Porém, ainda nos resta realizar a mudança de base no caso matrizes não diagonalizáveis. No caso diagonalizável, sabemos que a matriz de mudança de base é dada por suas colunas preenchidas com os autovetores da matriz, mas no caso não diagonalizável não conseguimos achar um conjunto de autovetores da mesma maneira.

Como qualquer matriz \mathbb{M} deve ser similar à uma matriz \mathbb{J} na forma de Jordan, deve existir uma matriz \mathbb{P} invertível tal que $\mathbb{M} = \mathbb{P}\mathbb{J}\mathbb{P}^{-1}$, e podemos, com isto, calcular as potências $\mathbb{M}^n = \mathbb{P}\mathbb{J}^n\mathbb{P}^{-1}$. Porém, o processo para encontrar a matriz \mathbb{P} no caso de matrizes não diagonalizáveis é um pouco mais trabalhoso, e dado que, caso queiramos estudar um caso prático, é possível calcular a forma de Jordan de qualquer matriz facilmente com o auxílio de um computador, não entraremos em detalhes do processo. Mas, caso o leitor tenha interesse, o capítulo 9 de (BRONSON, 1991) explica muito claramente cada passo dos cálculos necessários.

2 O Caso Unidimensional

Primeiramente, vamos avaliar o sistema unidimensional, pela sua simplicidade, para mostrar como o mapa pode apresentar uma distribuição limite não gaussiana, e com difusão não normal.

No caso unidimensional, nosso sistema tem a forma:

$$x[n+1] = Ax[n] + \eta[n+1], \quad n \in \{0, 1, 2, \dots, \infty\}, \quad (2.1)$$

onde o $x[n]$ é um escalar, assim como A , e η_i é uma variável aleatória com função distribuição $f_i(\eta_i)$. Note que o subíndice em f_i indica que a função distribuição do ruído η pode variar ao longo da evolução do sistema. Vamos considerar o caso em que o ruído tem sua distribuição centralizada, $\langle \eta[j] \rangle = 0 \forall j$, mas não apresenta correlação temporal, ou seja, $\langle \eta_i \eta_j \rangle = \langle \eta_i \rangle \langle \eta_j \rangle$, $i \neq j$.

Podemos escrever o n -ésimo termo da soma através do esquema triangular de Kolmogorov ([GNEDENKO; KOLMOGOROV, 1968](#)):

$$x[1] = Ax[0] + \eta[1]; \quad (2.2)$$

$$x[2] = A^2x[0] + A\eta[1] + \eta[2]; \quad (2.3)$$

$$x[3] = A^3x[0] + A^2\eta[1] + A\eta[2] + \eta[3]; \quad (2.4)$$

⋮

$$x[n] = A^n x[0] + S[n], \quad S[n] = \sum_{j=1}^n A^{n-j} \eta[j]. \quad (2.5)$$

Se a função característica de $\eta[j]$ é dada por

$$\psi_{\eta[j]}(z) = \eta_j(z) = \exp \left[-\frac{1}{2} m_j^2 z^2 (1 + \omega_j(z)) \right], \quad (2.6)$$

sendo m_j^2 a variância de $\eta[j]$, podemos escrever a função característica de $S[n]$:

$$\begin{aligned} \psi_{S[n]}(z) &= \left\langle \exp \left(Iz \sum_{j=1}^n A^{n-j} \eta[j] \right) \right\rangle = \prod_{j=1}^n \left\langle \exp \left(Iz A^{n-j} \eta[j] \right) \right\rangle \\ &= \prod_{j=1}^n \exp \left[-\frac{1}{2} A^{2(n-j)} m_j^2 z^2 (1 + \omega_j(A^{n-j} z)) \right] \\ &= \exp \left\{ -\frac{1}{2} M_n^2 z^2 [1 + \omega_{S[n]}(z)] \right\}, \end{aligned} \quad (2.7)$$

$$M_n^2 = \sum_{j=1}^n A^{2(n-j)} m_j^2; \quad (2.8)$$

$$\omega_{S[n]}(z) = \frac{1}{M_n^2} \sum_{j=1}^n A^{2(n-j)} \omega_j(A^{n-j} z). \quad (2.9)$$

Deste modo, a distribuição limite de $S[n]$ será gaussiana somente se $A^{2(n-j)}/M_n^2 \rightarrow 0 \forall j$, à medida que $n \rightarrow \infty$, anulando a contribuição das funções ω_j . Porém, os cálculos feitos até aqui são muito gerais para determinar qualquer propriedade da distribuição limite do mapa. Na próxima seção, vamos avaliar o caso em que o ruído é estacionário, ou seja, sua distribuição é a mesma durante todo o tempo. Isso simplifica a análise e permite que tiremos conclusões sobre a distribuição limite do sistema baseado no fator A do sistema.

2.1 Ruído Estacionário

Vamos apresentar aqui um caso especial, onde o ruído $\eta[j]$ é estacionário, ou seja, o ruído tem a mesma função distribuição ao longo de toda a evolução do sistema, $f_j(\eta[j]) = f(\eta) \forall j$. Será dado foco a este caso ao longo de todo o texto, dada sua simplicidade. Como será visto, este caso extremamente simples é um exemplo de não aplicabilidade do teorema do limite central. Novamente, consideramos que η é centralizado, não apresenta correlação temporal e sua variância é $\langle \eta[j]^2 \rangle = m^2 \forall j$. Note que, dada a estacionariedade do ruído, podemos escrever:

$$S[n] = \sum_{j=1}^n A^{n-j} \eta[j] = \sum_{j=0}^{n-1} A^j \eta[j]. \quad (2.10)$$

Neste caso, temos:

$$M_n^2 = m^2 \sum_{j=1}^{n-1} A^{2j} = m^2 \frac{A^{2n} - 1}{A^2 - 1}. \quad (2.11)$$

Com isto, podemos avaliar a gaussianidade da distribuição limite, assim como a dinâmica de difusão do sistema, em termos do módulo de A . Supondo $|A| < 1$, sabemos que a razão mA^j/M_n é máximo para $j = 0$:

$$\frac{m^2}{M_n^2} = \frac{1 - A^2}{1 - A^{2n}}, \quad (2.12)$$

que é diferente de zero. No limite da soma infinita temos:

$$\lim_{n \rightarrow \infty} \frac{m^2}{M_n^2} = 1 - A^2. \quad (2.13)$$

Portanto, para $|A| < 1$, a função ω não será anulada, ao menos para $j = 0$, de modo que a distribuição limite não será gaussiana. Além disso, note que o sistema não apresenta difusão normal. Mais do que subdifusão, o mapa apresenta saturação na variância, com esta convergindo para $m^2/(1 - A^2)$. Agora, se $|A| > 1$, A^j terá seu maior valor para $j = n - 1$ e a razão será dada por:

$$\frac{m^2 A^{2(n-1)}}{M_n^2} = A^{2(n-1)} \frac{A^2 - 1}{A^{2n} - 1} = \frac{1 - A^{-2}}{1 - A^{-2n}}. \quad (2.14)$$

$$\therefore \lim_{n \rightarrow \infty} \frac{m A^{n-1}}{M_n} = \sqrt{1 - A^{-2}}. \quad (2.15)$$

Concluimos que, também para $|A| > 1$, a distribuição limite também não deve ser gaussiana, mas, desta vez teremos um caso de difusão geométrica, com a variância crescendo exponencialmente de acordo com A^{2n} . Porém, se $|A| = 1$ temos, para todo j :

$$M_n^2 = nm^2 \Rightarrow \lim_{n \rightarrow \infty} \frac{m^2}{M_n^2} = 0 \quad (2.16)$$

Portanto, quando $|A| = 1$, a distribuição do sistema eventualmente convergirá para uma gaussiana, e apresentará difusão normal, com a variância proporcional a n . Mas, caso $A \neq 1$, a distribuição da soma dependerá da função $\omega(z)$, e portanto não terá uma distribuição gaussiana – a menos, é claro, que cada variável aleatória da soma tenha distribuição gaussiana, que leva a $\omega(z) = 0 \forall z$.

Note que, como foi discutido na introdução, A deve ser da forma $A = e^{A'\Delta t}$, sendo Δt o intervalo de tempo durante o qual o ruído é constante. Portanto, se tomamos o limite $\Delta t \rightarrow 0$, com o objetivo de recuperar o comportamento esperado para a equação diferencial estocástica, temos $A \rightarrow 1$, o que significa que o único comportamento que podemos esperar é aquele da distribuição limite gaussiana, com difusão normal. Nesse aspecto, vemos que a dinâmica do mapa discreto pode ser mais rica que aquela da equação diferencial. Porém é importante lembrar que a simplicidade da dinâmica da equação diferencial estocástica neste caso está associada ao fato de ser uma equação linear. De fato, é possível elaborar dinâmicas tão complexas quanto se queira, utilizando equações não lineares e ruído multiplicativo, dependente da posição no espaço de fase.

3 Soluções Gerais para a Função Característica do Mapa

Neste capítulo, desejamos apresentar a existência de fórmulas gerais para a função característica da distribuição do sistema

$$\mathbf{x}[n + 1] = \mathbb{A}\mathbf{x}[n] + \eta[n + 1], \quad n = 0, 1, 2, \dots, N, \quad (3.1)$$

sendo \mathbf{x} um vetor p -dimensional, assim como o vetor aleatório η , e \mathbb{A} é uma matriz $p \times p$. Veremos que, apesar da existência destas fórmulas, elas não são tão convenientes para uma avaliação da distribuição limite do sistema, dada sua grande complexidade.

Aqui, consideramos o sistema com algumas condições sobre o ruído η :

1. $\eta[n]$ é centralizado, $\langle \eta[n] \rangle = 0$ e sua matriz de covariância \mathfrak{m} é finita, lembrando que $\mathfrak{m}_{i,j} = \langle \eta_i \eta_j \rangle$ para um vetor centralizado.
2. η é um vetor aleatório estacionário, ou seja, sua distribuição de probabilidades é a mesma para qualquer instante de tempo.
3. Não há dependência entre os valores de η em instantes de tempo diferentes, ou seja, $\eta[n]$ e $\eta[m]$ são independentes para todo $n \neq m$.

Note que não assumimos independência entre as componentes do ruído, o que significa que, mesmo na base do seu vetor principal, por mais que as componentes do ruído não apresentem correlações lineares, elas ainda podem ser dependentes entre si, apresentando correlações de maiores ordens.

Primeiramente, vamos avaliar o mapa em termos do vetor principal do ruído. Com isso, obtemos uma fórmula geral para a função característica de um sistema qualquer, em termos dos autovalores da matriz de covariância (ou de correlação) do ruído η , e das matrizes de mudança de base para sua diagonalização (transformação necessária para obtenção do vetor principal).

Em seguida, na tentativa de simplificarmos a fórmula final, consideramos o mapa na base que diagonaliza a matriz \mathbb{A} , de modo que cada componente do sistema possa ser tratada independentemente como um sistema unidimensional. Apesar de conseguirmos resultados promissores para cada componente, surge uma dificuldade em combinar as componentes para obtenção da função característica conjunta do sistema considerando

todas as suas componentes. A partir disso, concluímos que, se desejamos avaliar o sistema componente a componente, para posteriormente recuperarmos a função característica conjunta, precisamos simplificar o sistema, assumindo que o vetor principal do ruído tem suas componentes independentes entre si, mesmo que nesta base a matriz \mathbb{A} não seja diagonalizada. No Capítulo 4, consideramos esta análise e determinamos condições necessárias para a gaussianização da distribuição limite do mapa.

Por fim, considerando a possibilidade de que a matriz \mathbb{A} não seja diagonalizável, estudamos o mapa na base que leva \mathbb{A} à sua forma canônica de Jordan.

3.1 Solução na Base do Vetor Principal para o Ruído

Primeiramente, vamos passar o problema para a base do vetor principal $\eta^{(p)} = \mathbb{T}\eta$, de modo que

$$\mathbb{T}\mathbb{m}\bar{\mathbb{T}} = \mathbb{D}_m^2 = \text{diag} \left(m_1^2, m_2^2, \dots, m_p^2 \right). \quad (3.2)$$

Como já sabemos, a matriz \mathbb{T} é ortogonal, o que implica $\mathbb{T}\bar{\mathbb{T}} = \mathbb{1}$, onde a barra denota transposição (e conjugação, no caso de matrizes complexas). Vamos aplicar essa transformação na equação do mapa:

$$\mathbb{T}\mathbf{x}[n+1] = \mathbb{T}\mathbb{A}\bar{\mathbb{T}}\mathbb{T}\mathbf{x}[n] + \mathbb{T}\eta[n+1] \quad (3.3)$$

$$\mathbf{y}[n+1] = \mathbb{B}\mathbf{y}[n] + \eta^{(p)}[n+1], \quad (3.4)$$

ou seja, $\mathbf{y}[n] = \mathbb{T}\mathbf{x}[n]$ e $\mathbb{B} = \mathbb{T}\mathbb{A}\bar{\mathbb{T}}$.

Lembremos que o vetor principal normalizado $\bar{\eta}^{(p)}$ é tal que $\eta^{(p)} = \mathbb{D}_m\bar{\eta}^{(p)}$, e que a função característica do vetor principal normalizado é dada por

$$\psi_{\bar{\eta}^{(p)}}(\mathbf{z}) = \exp \left[-\frac{\mathbf{z} \cdot \mathbf{z}}{2} \left(1 + \omega_{\bar{\eta}^{(p)}}(\mathbf{z}) \right) \right]. \quad (3.5)$$

Deste modo:

$$\begin{aligned} \psi_{\eta^{(p)}}(\mathbf{z}) &= \left\langle e^{I\mathbf{z} \cdot \eta^{(p)}} \right\rangle = \left\langle e^{I\mathbf{z} \cdot \mathbb{D}_m\bar{\eta}^{(p)}} \right\rangle = \left\langle e^{I(\mathbb{D}_m\mathbf{z}) \cdot \bar{\eta}^{(p)}} \right\rangle = \psi_{\bar{\eta}^{(p)}}(\mathbb{D}_m\mathbf{z}) \\ &= \exp \left\{ -\frac{1}{2} (\mathbb{D}_m\mathbf{z}) \cdot (\mathbb{D}_m\mathbf{z}) \left[1 + \omega_{\bar{\eta}^{(p)}}(\mathbb{D}_m\mathbf{z}) \right] \right\} \\ &= \exp \left\{ -\frac{1}{2} \mathbf{z} \mathbb{D}_m^2 \mathbf{z} \left[1 + \omega_{\bar{\eta}^{(p)}}(\mathbb{D}_m\mathbf{z}) \right] \right\}. \end{aligned} \quad (3.6)$$

Com isto, podemos passar para o cálculo da função característica de $\mathbf{y}[n]$. Note que, a partir disso, podemos calcular a função característica de $\mathbf{x}[n]$:

$$\psi_{\mathbf{x}[n]}(\mathbf{z}) = \langle e^{I\mathbf{z} \cdot \mathbf{x}[n]} \rangle = \langle e^{I\mathbf{z} \cdot \mathbb{T}\mathbf{y}[n]} \rangle = \langle e^{I(\mathbb{T}\mathbf{z}) \cdot \mathbf{y}[n]} \rangle = \psi_{\mathbf{y}[n]}(\mathbb{T}\mathbf{z}). \quad (3.7)$$

Podemos escrever $\mathbf{y}[n]$ como uma soma de variáveis aleatórias:

$$\mathbf{y}[n] = \mathbb{B}^n \mathbf{y}[0] + \mathbf{S}[n], \quad \mathbf{S}[n] = \sum_{j=1}^n \mathbb{B}^{n-j} \eta^{(p)}[j]. \quad (3.8)$$

Como $\eta^{(p)}$ é estacionário, podemos escrever, de forma mais conveniente:

$$\mathbf{S}[n] = \sum_{j=0}^{n-1} \mathbb{B}^j \eta^{(p)}[j]. \quad (3.9)$$

Vamos calcular a função característica de $\mathbf{S}[n]$:

$$\begin{aligned} \psi_{\mathbf{S}[n]}(\mathbf{z}) &= \langle \exp \{ I\mathbf{z} \cdot \mathbf{S}[n] \} \rangle = \left\langle \exp \left\{ I\mathbf{z} \cdot \sum_{j=0}^{n-1} \mathbb{B}^j \eta^{(p)}[j] \right\} \right\rangle \\ &= \left\langle \exp \left\{ I \sum_{j=0}^{n-1} \mathbf{z} \cdot \mathbb{B}^j \eta^{(p)}[j] \right\} \right\rangle = \prod_{j=0}^{n-1} \langle \exp \{ I\mathbf{z} \cdot \mathbb{B}^j \eta^{(p)}[j] \} \rangle \\ &= \prod_{j=0}^{n-1} \langle \exp \{ I(\bar{\mathbb{B}}^j \mathbf{z}) \cdot \eta^{(p)}[j] \} \rangle = \prod_{j=0}^{n-1} \psi_{\eta^{(p)}}(\bar{\mathbb{B}}^j \mathbf{z}) \\ &= \prod_{j=0}^{n-1} \exp \left\{ -\frac{1}{2} \bar{\mathbb{B}}^j \mathbf{z} \mathbb{D}_m^2 \bar{\mathbb{B}}^j \mathbf{z} \left[1 + \omega_{\bar{\eta}^{(p)}}(\mathbb{D}_m \bar{\mathbb{B}}^j \mathbf{z}) \right] \right\} \\ &= \prod_{j=0}^{n-1} \exp \left\{ -\frac{1}{2} \mathbf{z} \mathbb{B}^j \mathbb{D}_m^2 \bar{\mathbb{B}}^j \mathbf{z} \left[1 + \omega_{\bar{\eta}^{(p)}}(\mathbb{D}_m \bar{\mathbb{B}}^j \mathbf{z}) \right] \right\} \\ &= \exp \left\{ -\frac{1}{2} \sum_{j=0}^{n-1} \mathbf{z} \mathbb{B}^j \mathbb{D}_m^2 \bar{\mathbb{B}}^j \mathbf{z} \left[1 + \omega_{\bar{\eta}^{(p)}}(\mathbb{D}_m \bar{\mathbb{B}}^j \mathbf{z}) \right] \right\}. \end{aligned} \quad (3.10)$$

Note que $\mathbb{B}^j \mathbb{D}_m^2 \bar{\mathbb{B}}^j$ é uma matriz simétrica:

$$\begin{aligned} \left[\mathbb{B}^j \mathbb{D}_m^2 \bar{\mathbb{B}}^j \right]_{i,k} &= \sum_{r,s} \mathbb{B}_{i,r}^j \left(\mathbb{D}_m^2 \right)_{r,s} \bar{\mathbb{B}}_{s,k}^j = \sum_{r,s} \mathbb{B}_{i,r}^j m_r^2 \delta_{r,s} \bar{\mathbb{B}}_{s,k}^j \\ &= \sum_r \mathbb{B}_{i,r}^j m_r^2 \bar{\mathbb{B}}_{r,k}^j = \sum_r \bar{\mathbb{B}}_{r,i}^j m_r^2 \mathbb{B}_{k,r}^j \\ &= \sum_r \mathbb{B}_{k,r}^j m_r^2 \bar{\mathbb{B}}_{r,i}^j \\ &= \left[\mathbb{B}^j \mathbb{D}_m^2 \bar{\mathbb{B}}^j \right]_{k,i}, \end{aligned} \quad (3.11)$$

onde utilizamos o fato de que a potência da transposta de uma matriz é igual a transposta da potência da matriz. Sendo assim,

$$\mathbb{M}[n] = \sum_{j=0}^{n-1} \mathbb{B}^j \mathbb{D}_m \bar{\mathbb{B}}^j \quad (3.12)$$

deve ser uma matriz simétrica, e podemos escrever:

$$\psi_{\mathbf{S}[n]}(\mathbf{z}) = \exp \left\{ -\frac{1}{2} \mathbf{z} \mathbb{M}[n] \mathbf{z} + \sum_{j=0}^{n-1} \left(\mathbf{z} \mathbb{B}^j \mathbb{D}_m^2 \bar{\mathbb{B}}^j \mathbf{z} \right) \omega_{\bar{\eta}^{(p)}} \left(\mathbb{D}_m \bar{\mathbb{B}}^j \mathbf{z} \right) \right\}. \quad (3.13)$$

Observando esta última forma, fica claro que $\mathbb{M}[n]$ é a matriz de covariância de $\mathbf{S}[n]$, $M_{i,j}[n] = \langle S_i[n] S_j[n] \rangle$. Dada a simetria de $\mathbb{M}[n]$, ela pode ser diagonalizada, através de uma transformação descrita por uma matriz ortogonal $\mathbb{P}[n]$, tal que $\mathbb{P}[n] \bar{\mathbb{P}}[n] = \bar{\mathbb{P}}[n] \mathbb{P}[n] = \mathbb{1}$, de modo que $\mathbb{P} \mathbb{M}[n] \bar{\mathbb{P}} = \mathbb{D}_{\mathbb{M}[n]}^2$, sendo $\mathbb{D}_{\mathbb{M}[n]} = \text{diag}(M_1[n], M_2[n], \dots, M_p[n])$. Com isto, calculamos o vetor principal $\mathbf{S}^{(p)}[n] = \mathbb{P}[n] \mathbf{S}[n]$, e sua função característica é dada por:

$$\psi_{\mathbf{S}^{(p)}[n]}(\mathbf{z}) = \langle e^{I \mathbf{z} \cdot \mathbf{S}^{(p)}[n]} \rangle = \langle e^{I \mathbf{z} \cdot (\mathbb{P}[n] \mathbf{S}[n])} \rangle = \langle e^{I (\bar{\mathbb{P}}[n] \mathbf{z}) \cdot \mathbf{S}[n]} \rangle = \psi_{\mathbf{S}[n]}(\bar{\mathbb{P}}[n] \mathbf{z}) \quad (3.14)$$

$$\begin{aligned} \therefore \psi_{\mathbf{S}^{(p)}[n]}(\mathbf{z}) &= \exp \left\{ -\frac{1}{2} \bar{\mathbb{P}}[n] \mathbf{z} \mathbb{M}[n] \bar{\mathbb{P}}[n] \mathbf{z} + \right. \\ &\quad \left. + \sum_{j=0}^{n-1} \left(\bar{\mathbb{P}}[n] \mathbf{z} \cdot \mathbb{B}^j \mathbb{D}_m^2 \bar{\mathbb{B}}^j \bar{\mathbb{P}}[n] \mathbf{z} \right) \omega_{\bar{\eta}^{(p)}} \left(\mathbb{D}_m \bar{\mathbb{B}}^j \bar{\mathbb{P}}[n] \mathbf{z} \right) \right\} \\ &= \exp \left\{ -\frac{1}{2} \mathbf{z} \left(\mathbb{P}[n] \mathbb{M}[n] \bar{\mathbb{P}}[n] \right) \mathbf{z} + \right. \\ &\quad \left. + \sum_{j=0}^{n-1} \left(\mathbf{z} \mathbb{P}[n] \mathbb{B}^j \mathbb{D}_m^2 \bar{\mathbb{B}}^j \bar{\mathbb{P}}[n] \mathbf{z} \right) \omega_{\bar{\eta}^{(p)}} \left(\mathbb{D}_m \bar{\mathbb{B}}^j \bar{\mathbb{P}}[n] \mathbf{z} \right) \right\} \\ &= \exp \left\{ -\frac{1}{2} \mathbf{z} \mathbb{D}_{\mathbb{M}[n]}^2 \mathbf{z} + \right. \\ &\quad \left. + \sum_{j=0}^{n-1} \left(\mathbf{z} \mathbb{P}[n] \mathbb{B}^j \mathbb{D}_m^2 \bar{\mathbb{B}}^j \bar{\mathbb{P}}[n] \mathbf{z} \right) \omega_{\bar{\eta}^{(p)}} \left(\mathbb{D}_m \bar{\mathbb{B}}^j \bar{\mathbb{P}}[n] \mathbf{z} \right) \right\}. \quad (3.15) \end{aligned}$$

Passando agora para o vetor principal reduzido (ou normalizado) de $\mathbf{S}[n]$, sabemos que $\mathbf{S}[n] = \mathbb{D}_M \bar{\mathbf{S}}[n] \Rightarrow \bar{\mathbf{S}}[n] = \mathbb{D}_M^{-1} \mathbf{S}[n]$, de modo que

$$\psi_{\bar{\mathbf{S}}[n]}(\mathbf{z}) = \langle e^{I \mathbf{z} \cdot \bar{\mathbf{S}}[n]} \rangle = \langle e^{I \mathbf{z} \cdot (\mathbb{D}_M^{-1} \mathbf{S}[n])} \rangle = \langle e^{I (\mathbb{D}_M^{-1} \mathbf{z}) \cdot \mathbf{S}[n]} \rangle = \psi_{\mathbf{S}[n]}(\mathbb{D}_M^{-1} \mathbf{z}). \quad (3.16)$$

A função característica do vetor principal reduzido deve ser dada por

$$\psi_{\bar{\mathbf{S}}[n]}(\mathbf{z}) = \exp \left\{ -\frac{\mathbf{z} \cdot \mathbf{z}}{2} \left[1 + \omega_{\bar{\mathbf{S}}[n]}(\mathbf{z}) \right] \right\}, \quad (3.17)$$

sendo

$$\omega_{\bar{\mathbf{S}}[n]}(\mathbf{z}) = \frac{\sum_{j=0}^{n-1} \left(\mathbf{z} \mathbb{D}_{\mathbb{M}[n]}^{-1} \mathbb{P}[n] \mathbb{B}^j \mathbb{D}_{\mathbb{m}}^2 \bar{\mathbb{B}}^j \bar{\mathbb{P}}[n] \mathbb{D}_{\mathbb{M}[n]}^{-1} \mathbf{z} \right) \omega_{\bar{\eta}(p)} \left(\mathbb{D}_{\mathbb{m}} \bar{\mathbb{B}}^j \bar{\mathbb{P}}[n] \mathbb{D}_{\mathbb{M}[n]}^{-1} \mathbf{z} \right)}{\mathbf{z} \cdot \mathbf{z}}. \quad (3.18)$$

Concluimos que somos capazes de expressar a função característica da distribuição solução do mapa com ruído aditivo. Porém, considerando apenas a forma geral obtida para a função característica, não conseguimos extrair informações relevantes sobre a distribuição do sistema no limite $n \rightarrow \infty$. Na próxima seção, vamos analisar o sistema na base que diagonaliza a matriz \mathbb{A} , caso esta base exista.

3.2 Solução na Base que Diagonaliza o Mapa

Na seção prévia, nos preocupamos apenas em diagonalizar a matriz de covariância do ruído, para escrever os resultados em termos de seus autovalores. Agora, consideremos o mapa p -dimensional

$$\mathbf{x}[n+1] = \mathbb{A}\mathbf{x}[n] + \eta[n+1] \quad (3.19)$$

com ruído η centralizado. Vamos supor que \mathbb{A} possa ser diagonalizada, de modo que $\mathbb{L}\mathbb{A}\mathbb{L}^{-1} = \mathbb{D}_{\mathbb{A}} = \text{diag}(A_1, A_2, \dots, A_p)$. Assim como na seção anterior, supomos que $\eta[n]$ e $\eta[m]$ são independentes para quaisquer $n \neq m$, e supomos que o ruído é estacionário, ou seja, a função distribuição de $\eta[n]$ é a mesma para qualquer valor de n . Note que não fazemos suposições sobre as relações entre as componentes de η , de modo que estas podem apresentar dependências entre si.

Definindo $\mathbf{y} = \mathbb{L}\mathbf{x} \Rightarrow \mathbf{x} = \mathbb{L}^{-1}\mathbf{y}$ e $\zeta = \mathbb{L}\eta$, podemos reescrever o mapa:

$$\mathbb{L}\mathbf{x}[n+1] = \mathbb{L}\mathbb{A}\mathbb{L}^{-1}\mathbb{L}\mathbf{x}[n] + \mathbb{L}\eta[n+1] \quad (3.20)$$

$$\mathbf{y}[n+1] = \mathbb{D}_{\mathbb{A}}\mathbf{y}[n] + \zeta[n+1]. \quad (3.21)$$

Como vimos na seção passada, dado que \mathbf{x} e \mathbf{y} estão relacionados por transformações lineares, é simples relacionar suas funções características. Isto é necessário pois, se \mathbf{x} é

o vetor de nosso interesse, é necessário que possamos recuperar sua função distribuição após a consideração do sistema em outra base. A relação entre as funções características é dada por:

$$\psi_{\mathbf{x}[n]}(\mathbf{z}) = \langle e^{I\mathbf{z} \cdot \mathbf{x}[n]} \rangle = \langle e^{I\mathbf{z} \cdot (\mathbb{L}^{-1}\mathbf{y}[n])} \rangle = \langle e^{I(\bar{\mathbb{L}}^{-1}\mathbf{z}) \cdot \mathbf{y}[n]} \rangle \quad (3.22)$$

$$\therefore \psi_{\mathbf{x}[n]}(\mathbf{z}) = \psi_{\mathbf{y}[n]}(\bar{\mathbb{L}}^{-1}\mathbf{z}) \Rightarrow \psi_{\mathbf{y}[n]}(\mathbf{z}) = \psi_{\mathbf{x}[n]}(\bar{\mathbb{L}}\mathbf{z}). \quad (3.23)$$

Da mesma maneira, $\psi_{\zeta}(\mathbf{z}) = \psi_{\eta}(\bar{\mathbb{L}}\mathbf{z})$ (note que a função característica do ruído não deve depender do instante de tempo, dado que o ruído é estacionário). A análise a seguir é semelhante ao que foi feito na seção anterior. Note que podemos escrever:

$$\mathbf{y}[n] = \mathbb{D}_{\mathbb{A}}^n \mathbf{y}[0] + \mathbf{S}[n], \quad \mathbf{S}[n] = \sum_{j=1}^n \mathbb{D}_{\mathbb{A}}^{n-j} \zeta[j] = \sum_{j=0}^{n-1} \mathbb{D}_{\mathbb{A}}^j \zeta[j], \quad (3.24)$$

onde a última igualdade é válida devido à estacionariedade do ruído. A função característica de $\mathbf{S}[n]$ é dada por

$$\psi_{\mathbf{S}[n]}(\mathbf{z}) = \left\langle \exp \left\{ I\mathbf{z} \cdot \sum_{j=0}^{n-1} \mathbb{D}_{\mathbb{A}}^j \zeta[j] \right\} \right\rangle = \prod_{j=0}^{n-1} \left\langle \exp \left\{ I\mathbf{z} \cdot \mathbb{D}_{\mathbb{A}}^j \zeta[j] \right\} \right\rangle = \prod_{j=0}^{n-1} \psi_{\zeta}(\mathbb{D}_{\mathbb{A}}^j \mathbf{z}), \quad (3.25)$$

onde a última igualdade é obtida lembrando que a transposta de uma matriz diagonal é igual à matriz original. Lembrando que o ruído η é centralizado, ζ também deve ser, o que implica que sua função característica deve poder ser escrita na forma:

$$\psi_{\zeta}(\mathbf{z}) = \psi_{\eta}(\mathbf{z}) = \exp \left[-\frac{1}{2} \mathbf{z} \mathbb{M} \mathbf{z} (1 + \omega_{\zeta}(\mathbf{z})) \right], \quad (3.26)$$

sendo $\mathbb{M}_{i,j} = \langle \zeta_i \zeta_j \rangle$ a matriz de covariância do vetor aleatório ζ . Vamos explicitar a relação entre \mathbb{M} e a matriz de covariância \mathbb{N} , $\mathbb{N}_{i,j} = \langle \eta_i \eta_j \rangle$, do ruído original η :

$$\begin{aligned} \mathbb{M}_{i,j} &= \langle \zeta_i \zeta_j \rangle = \langle (\mathbb{L}\eta)_i (\mathbb{L}\eta)_j \rangle = \left\langle \left(\sum_{k=1}^p \mathbb{L}_{i,k} \eta_k \right) \left(\sum_{l=1}^p \mathbb{L}_{j,l} \eta_l \right) \right\rangle \\ &= \sum_{k,l} \langle \mathbb{L}_{i,k} \mathbb{L}_{j,l} \eta_k \eta_l \rangle = \sum_{k,l} \mathbb{L}_{i,k} \mathbb{L}_{j,l} \langle \eta_k \eta_l \rangle = \sum_{k,l} \mathbb{L}_{i,k} \mathbb{L}_{j,l} \mathbb{N}_{k,l} \\ &= \sum_{k,l} \mathbb{L}_{i,k} \mathbb{N}_{k,l} \bar{\mathbb{L}}_{l,j} = \left(\mathbb{L} \mathbb{N} \bar{\mathbb{L}} \right)_{i,j} \end{aligned} \quad (3.27)$$

$$\therefore \mathbb{M} = \mathbb{L} \mathbb{N} \bar{\mathbb{L}}. \quad (3.28)$$

Voltando à função característica de $\mathbf{S}[n]$:

$$\begin{aligned}\psi_{\mathbf{S}[n]}(\mathbf{z}) &= \prod_{j=0}^{n-1} \exp \left\{ -\frac{1}{2} \left(\mathbb{D}_{\mathbb{A}}^j \mathbf{z} \right) \cdot \mathbb{M} \left(\mathbb{D}_{\mathbb{A}}^j \mathbf{z} \right) \left[1 + \omega_{\zeta} \left(\mathbb{D}_{\mathbb{A}}^j \mathbf{z} \right) \right] \right\} \\ &= \prod_{j=0}^{n-1} \exp \left\{ -\frac{1}{2} \mathbf{z} \left(\mathbb{D}_{\mathbb{A}}^j \mathbb{M} \mathbb{D}_{\mathbb{A}}^j \right) \mathbf{z} \left[1 + \omega_{\zeta} \left(\mathbb{D}_{\mathbb{A}}^j \mathbf{z} \right) \right] \right\}.\end{aligned}\quad (3.29)$$

Definindo $\mathbb{M}^{(n)} = \sum_{j=0}^{n-1} \mathbb{D}_{\mathbb{A}}^j \mathbb{M} \mathbb{D}_{\mathbb{A}}^j$, escrevemos:

$$\psi_{\mathbf{S}[n]}(\mathbf{z}) = \exp \left[-\frac{1}{2} \mathbf{z} \mathbb{M}^{(n)} \mathbf{z} \left(1 + \omega_{\mathbf{S}[n]}(\mathbf{z}) \right) \right], \quad (3.30)$$

$$\omega_{\mathbf{S}[n]}(\mathbf{z}) = \frac{\sum_{j=0}^{n-1} \mathbf{z} \left(\mathbb{D}_{\mathbb{A}}^j \mathbb{M} \mathbb{D}_{\mathbb{A}}^j \right) \mathbf{z} \omega_{\zeta} \left(\mathbb{D}_{\mathbb{A}}^j \mathbf{z} \right)}{\mathbf{z} \mathbb{M}^{(n)} \mathbf{z}}. \quad (3.31)$$

Com isto, é claro que a matriz de covariância de $\mathbf{S}[n]$ é dada por $\mathbb{M}^{(n)} = \sum_{j=0}^{n-1} \mathbb{D}_{\mathbb{A}}^j \mathbb{M} \mathbb{D}_{\mathbb{A}}^j$, lembrando que \mathbb{M} é a matriz de covariância do ruído ζ .

Note que é possível obter as funções características marginais, para componentes específicas do sistema:

$$\begin{aligned}\psi_{S_i[n]}(z) &= \langle \exp (I z S_i[n]) \rangle \\ &= \langle \exp (I (z_1 = 0, z_2 = 0, \dots, z_i = z, \dots, z_p = 0) \cdot \mathbf{S}[n]) \rangle \\ &= \psi_{\mathbf{S}[n]}(z_1 = 0, z_2 = 0, \dots, z_i = z, \dots, z_p = 0).\end{aligned}\quad (3.32)$$

Vejamos como a escolha $\mathbf{z} = \mathbf{q} \equiv (z_1 = 0, z_2 = 0, \dots, z_i = z, \dots, z_p = 0)$ afeta os termos da função característica de $\mathbf{S}[n]$:

$$\mathbf{q} \mathbb{M}^{(n)} \mathbf{q} = \sum_{k,l} q_k \mathbb{M}_{k,l}^{(n)} q_l = \sum_{k,l} \delta_{i,k} z \mathbb{M}_{k,l}^{(n)} \delta_{i,l} z = \mathbb{M}_{i,i}^{(n)} z^2; \quad (3.33)$$

$$\begin{aligned}\mathbf{q} \left(\mathbb{D}_{\mathbb{A}}^j \mathbb{M} \mathbb{D}_{\mathbb{A}}^j \right) \mathbf{q} &= \sum_{k,l} q_k \left(\mathbb{D}_{\mathbb{A}}^j \mathbb{M} \mathbb{D}_{\mathbb{A}}^j \right)_{k,l} q_l = \sum_{k,k',l,l'} \delta_{i,k} z \left(\mathbb{D}_{\mathbb{A}}^j \right)_{k,k'} \mathbb{M}_{k',l'} \left(\mathbb{D}_{\mathbb{A}}^j \right)_{l',l} \delta_{i,l} z \\ &= \sum_{k,l} \left(\mathbb{D}_{\mathbb{A}}^j \right)_{i,k} \mathbb{M}_{k,l} \left(\mathbb{D}_{\mathbb{A}}^j \right)_{l,i} z^2 = \sum_{k,l} \mathbb{M}_{k,l} \delta_{i,k} A_k^j \delta_{l,i} A_l^j z^2 \\ &= \mathbb{M}_{i,i} A_i^{2j} z^2;\end{aligned}\quad (3.34)$$

$$\left(\mathbb{D}_{\mathbb{A}}^j \mathbf{q}\right)_l = \sum_{k=1}^p \left(\mathbb{D}_{\mathbb{A}}^j\right)_{l,k} q_k = \sum_{k=1}^p \delta_{l,k} A_k^j \delta_{i,k} z = \delta_{l,i} A_i^j z \quad (3.35)$$

$$\therefore \mathbb{D}_{\mathbb{A}}^j \mathbf{q} = \left(0, 0, \dots, 0, z_i = A_i^j z, 0, \dots, 0\right). \quad (3.36)$$

Note que para $\omega_{\zeta} \left(\mathbb{D}_{\mathbb{A}}^j \mathbf{q}\right) = \omega_{\zeta_i} \left(A_i^j z\right) = \omega_i \left(A_i^j z\right)$. Deste modo, a função $\omega_{\mathbf{S}[n]}$ é dada por

$$\omega_{\mathbf{S}[n]}(\mathbf{q}) = \frac{\sum_{j=0}^{n-1} A_i^{2j} \mathbb{M}_{i,i} z^2 \omega_i \left(A_i^j z\right)}{\mathbb{M}_{i,i}^{(n)} z^2} = \frac{\mathbb{M}_{i,i}}{\mathbb{M}_{i,i}^{(n)}} \sum_{j=0}^{n-1} A_i^{2j} \omega_i \left(A_i^j z\right), \quad (3.37)$$

sendo que podemos reescrever $\mathbb{M}_{i,i}^{(n)}$:

$$\begin{aligned} \mathbb{M}_{i,i}^{(n)} &= \sum_{j=0}^{n-1} \left(\mathbb{D}_{\mathbb{A}}^j \mathbb{M} \mathbb{D}_{\mathbb{A}}^j\right)_{i,i} = \sum_{j=0}^{n-1} \sum_{k,l} \left(\mathbb{D}_{\mathbb{A}}\right)_{i,k} \mathbb{M}_{k,l} \left(\mathbb{D}_{\mathbb{A}}\right)_{l,i} \\ &= \sum_{j=0}^{n-1} \sum_{k,l} \delta_{i,k} A_k^j \mathbb{M}_{k,l} \delta_{l,i} A_i^j = \sum_{j=0}^{n-1} \mathbb{M}_{i,i} A_i^{2j} = \mathbb{M}_{i,i} \sum_{j=0}^{n-1} A_i^{2j}. \end{aligned} \quad (3.38)$$

Portanto:

$$\begin{aligned} \psi_{S_i[n]}(z) &= \psi_{\mathbf{S}[n]}(\mathbf{q}) = \exp \left\{ -\frac{1}{2} \mathbb{M}_{i,i}^{(n)} z^2 \left[1 + \omega_{\mathbf{S}[n]}(\mathbf{q}) \right] \right\} \\ &= \exp \left\{ -\frac{1}{2} \mathbb{M}_{i,i} \left(\sum_{j=0}^{n-1} A_i^{2j} \right) z^2 \left[1 + \frac{\sum_{j=0}^{n-1} A_i^{2j} \omega_i \left(A_i^j z\right)}{\sum_{j=0}^{n-1} A_i^{2j}} \right] \right\}, \end{aligned} \quad (3.39)$$

lembrando que $\omega_i = \omega_{\zeta_i}$, ou seja, é a função ω do ruído na base que diagonaliza o sistema, e não na base do ruído original η . Perceba que o resultado obtido para uma única componente na base diagonalizada corresponde àquele apresentado no Capítulo 2, esperado de um mapa unidimensional, o que é de se esperar, uma vez que, ao diagonalizarmos o sistema, temos a evolução das componentes acontecendo de maneira independente. Com isso, podemos determinar se a distribuição limite nesta componente será ou não gaussiana, a depender do autovalor A_i .

Por um lado a expressão obtida para cada componente na base diagonalizada (3.39) é muito simples, facilitando a análise da distribuição limite das componentes do mapa diagonalizado. Note que, se formos capazes de encontrar condições para que os termos

envolvendo as funções ω_i se anularem, demonstramos a gaussianização da distribuição de S_i e, conseqüentemente, de y_i . Por outro lado, mesmo que demonstremos a gaussianização de cada componente y_i , não podemos afirmar que a distribuição de \mathbf{y} será uma gaussiana multidimensional, uma vez que, havendo correlações entre as componentes de ζ , não somos capazes de construir a função $\omega_{\mathbf{S}[n]}(\mathbf{z})$ a partir das funções $\omega_{S_i[n]}(z)$, $i \in \{1, 2, \dots, p\}$. Veja que

$$\psi_{\mathbf{S}[n]} = \langle e^{I\mathbf{z} \cdot \mathbf{S}[n]} \rangle = \left\langle \prod_{j=1}^p e^{Iz_j S_j[n]} \right\rangle \neq \prod_{j=1}^p \langle e^{Iz_j S_j[n]} \rangle = \prod_{j=1}^p \psi_{S_j[n]}, \quad (3.40)$$

onde a diferença aparece uma vez que $S_j[n]$ depende $\zeta_j[k] \forall k \in \{1, 2, \dots, n\}$ e podem haver dependências entre $S_i[k]$ e $S_j[k]$, $i \neq j$, de modo que as funções características não são separáveis.

Caso exista uma base na qual o vetor de ruído tem suas componentes independentes entre si, se conseguirmos demonstrar a gaussianização das componentes do sistema nessa base, demonstramos a gaussianização do sistema como um todo em qualquer base que seja. Dada a independência das componentes do ruído, a função característica deverá poder ser escrita como o produto das funções características de cada componentes, as quais são gaussianas. Assim, ao mudarmos de base, a distribuição da transformação de uma distribuição normal multivariada deve continuar sendo gaussiana, uma vez que cada componente na nova base será dada por uma combinação linear de distribuições gaussianas, que sabemos que deve também apresentar uma distribuição normal.

Note que, dado que a matriz de covariância sempre pode ser diagonalizada, e que seu conjunto de autovetores normalizados é único, a menos de uma normalização, a base na qual as componentes do ruído são independentes entre si deve ser aquela do vetor principal do ruído.

No próximo capítulo, vamos estudar o mapa na base do vetor principal do ruído, assumindo que este tem suas componentes independentes entre si, de modo a simplificar a análise do sistema, e vamos encontrar as condições necessárias para a gaussianização de cada componente. Com isso, podemos determinar quando o sistema como um todo converge para uma distribuição normal multivariada.

3.3 Solução para o Mapa Não Diagonalizável: Forma Canônica de Jordan

Caso a matriz \mathbb{A} não seja diagonalizável, o melhor que podemos fazer é passar para a base em que obtemos a forma canônica de Jordan da matriz: $\mathbb{L}\mathbb{A}\mathbb{L}^{-1} = \mathbb{J}_{\mathbb{A}} =$

$\text{diag}(\mathbb{J}_{(1)}, \mathbb{J}_{(2)}, \dots, \mathbb{J}_{(r)})$, sendo r a quantidade de blocos de Jordan, cada um com componentes $(\mathbb{J}_{(q)})_{i,j} = J_q \delta_{i,j} + \delta_{i,j-1}$, sendo J_q o autovalor do bloco.

Neste caso, a análise é muito semelhante àquela apresentada na seção anterior. Passamos da equação original para a base na qual temos a forma de Jordan de \mathbb{A} :

$$\mathbf{x}[n+1] = \mathbb{A}\mathbf{x}[n] + \eta[n+1] \Rightarrow \quad (3.41)$$

$$\mathbf{y}[n+1] = \mathbb{J}_{\mathbb{A}}\mathbf{y}[n] + \zeta[n+1], \quad (3.42)$$

sendo $\mathbf{y}[n] = \mathbb{L}\mathbf{x}[n]$ e $\zeta[n] = \mathbb{L}\eta[n]$. Novamente, assumimos que η é centralizado, estacionário e não apresenta correlação temporal. Repetindo os cálculos da sessão anterior, trocando a matriz diagonal pela forma de Jordan, escrevemos:

$$\mathbf{y}[n] = \mathbb{J}_{\mathbb{A}}^n \mathbf{y}[0] + \mathbf{S}[n], \quad \mathbf{S}[n] = \sum_{j=1}^n \mathbb{J}_{\mathbb{A}}^{n-j} \zeta[j] = \sum_{j=0}^{n-1} \mathbb{J}_{\mathbb{A}} \zeta[j]. \quad (3.43)$$

Se \mathbb{N} é a matriz de covariância do ruído original η , então $\mathbb{M} = \mathbb{L}\mathbb{N}\bar{\mathbb{L}}$ é a matriz de covariância de ζ . Com isto, escrevemos a função característica da soma $\mathbf{S}[n]$:

$$\psi_{\mathbf{S}[n]}(\mathbf{z}) = \exp \left[-\frac{1}{2} \mathbf{z} \mathbb{M}^{(n)} \mathbf{z} \left(1 + \omega_{\mathbf{S}[n]}(\mathbf{z}) \right) \right]; \quad (3.44)$$

$$\omega_{\mathbf{S}[n]}(\mathbf{z}) = \frac{\sum_{j=0}^{n-1} \mathbf{z} \left(\mathbb{J}_{\mathbb{A}}^j \mathbb{M} \bar{\mathbb{J}}_{\mathbb{A}}^j \right) \mathbf{z} \omega_{\zeta} \left(\bar{\mathbb{J}}_{\mathbb{A}}^j \mathbf{z} \right)}{\mathbf{z} \mathbb{M}^{(n)} \mathbf{z}}; \quad (3.45)$$

$$\mathbb{M}^{(n)} = \sum_{j=0}^{n-1} \mathbb{J}_{\mathbb{A}}^j \mathbb{M} \bar{\mathbb{J}}_{\mathbb{A}}^j. \quad (3.46)$$

Note que podemos calcular explicitamente os termos de $\mathbb{J}_{\mathbb{A}}^j \mathbb{M} \bar{\mathbb{J}}_{\mathbb{A}}^j$. Para isso, precisamos definir os termos da potência da matriz na forma de Jordan. Sabemos que esta deve ser dada pelas potências dos blocos de Jordan, ou seja:

$$(\mathbb{J}_{\mathbb{A}}^n)_{i,k} = \begin{cases} (\mathbb{J}_{(q)}^n)_{i-p'_{q-1}, k-p'_{q-1}}, & p'_{q-1} < i, k \leq p'_q; \\ 0, & \text{caso contrário,} \end{cases} \quad (3.47)$$

definindo:

$$p'_q = \sum_{i=0}^q p_i, \quad p_0 = 0 \quad (3.48)$$

e lembrando que os termos da potência de um bloco de Jordan são dados por

$$\left(\mathbb{J}_{(q)}^j\right)_{i,k} = \sum_{a=0}^{W_q} \binom{j}{a} J_q^{j-a} \delta_{i,k-a}, \quad W_q = \min \{p_q - 1, j\}. \quad (3.49)$$

Portanto, a transposta é dada por:

$$\left(\bar{\mathbb{J}}_{(q)}^j\right)_{i,k} = \sum_{a=0}^{W_q} \binom{j}{a} J_q^{j-a} \delta_{i,k+a} \quad (3.50)$$

Agora, podemos calcular de fato:

$$\begin{aligned} \left(\mathbb{J}_{\mathbb{A}}^j \mathbb{M} \bar{\mathbb{J}}_{\mathbb{A}}^j\right)_{i,k} &= \sum_{m,l} \left(\mathbb{J}_{\mathbb{A}}^j\right)_{i,m} \mathbb{M}_{m,l=1} \left(\bar{\mathbb{J}}_{\mathbb{A}}^j\right)_{l,k} \\ &= \sum_{q=1}^r \sum_{m,l}^{p_q} \left(\mathbb{J}_{\mathbb{A}}^j\right)_{i,m+p'_{q-1}} \mathbb{M}_{m+p'_{q-1},l+p'_{q-1}} \left(\bar{\mathbb{J}}_{\mathbb{A}}^j\right)_{l+p'_{q-1},k}. \end{aligned} \quad (3.51)$$

Note que, para $\left(\mathbb{J}_{\mathbb{A}}^j\right)_{i,m}$, precisamos considerar apenas os termos $i \leq m \leq p'_{q_i}$, sendo $q_i = \min \{q \in \{1, 2, \dots, r\}, p'_q \geq i\}$. Do mesmo modo, para a matriz transposta, precisamos considerar apenas os termos $k \leq l \leq p'_{q_k}$. Continuando com estas considerações:

$$\begin{aligned} \left(\mathbb{J}_{\mathbb{A}}^j \mathbb{M} \bar{\mathbb{J}}_{\mathbb{A}}^j\right)_{i,k} &= \sum_{m=i}^{p'_{q_i}} \sum_{l=k}^{p'_{q_k}} M_{m,l} \sum_{a_1=0}^{W_{q_i}} \binom{j}{a_1} J_{q_i}^{j-a_1} \delta_{i,m-a_1} \sum_{a_2=0}^{W_{q_i=k}} \binom{j}{a_2} J_{q_k}^{j-a_2} \delta_{l,k+a_2} \\ &= \sum_{m=i}^{p'_{q_i}} \sum_{l=k}^{p'_{q_k}} M_{m,l} \binom{j}{m-i} J_{q_i}^{j+i-m} \binom{j}{l-k} J_{q_k}^{j+k-l}. \end{aligned} \quad (3.52)$$

Podemos, ainda, escrever a função característica marginal para cada componente y_i , escolhendo $\mathbf{z} = \mathbf{z}' \equiv (z_1 = 0, z_2 = 0, \dots, z_i = z, \dots, z_p = 0)$, assim como fizemos no caso da matriz diagonalizada. Neste caso:

$$\mathbf{z}' \left(\mathbb{J}_{\mathbb{A}}^j \mathbb{M} \bar{\mathbb{J}}_{\mathbb{A}}^j\right) \mathbf{z}' = \left(\mathbb{J}_{\mathbb{A}}^j \mathbb{M} \bar{\mathbb{J}}_{\mathbb{A}}^j\right)_{i,i} z^2 = z^2 \sum_{m=i}^{p'_{q_i}} \sum_{l=i}^{p'_{q_i}} M_{m,l} \binom{j}{m-i} \binom{j}{l-i} J_{q_i}^{2(j+i)-m-l}; \quad (3.53)$$

$$\left(\bar{\mathbb{J}}_{\mathbb{A}}^j \mathbf{z}'\right)_l = \sum_{k=1}^p \left(\bar{\mathbb{J}}_{\mathbb{A}}\right)_{l,k} z_k = \left(\bar{\mathbb{J}}_{\mathbb{A}}\right)_{l,i} z = \begin{cases} z \sum_{a=0}^{W_{q_l}} \binom{j}{a} J_{q_l}^{j-a} \delta_{l,i+a}, & p_{q_{l-1}} < i \leq p_{q_l}, \\ 0, & \text{caso contrário.} \end{cases} \quad (3.54)$$

Com isto, estão explicitamente expressos os termos envolvidos na função característica de $S_i[n]$:

$$\psi_{S_i[n]}(z) = \psi_{\mathbf{S}[n]}(\mathbf{z}') = \exp \left[-\frac{1}{2} \left(\mathbb{M}^{(n)} \right)_{i,i} z^2 \left(1 + \omega_{S_i[n]}(z) \right) \right]; \quad (3.55)$$

$$\omega_{S_i[n]}(z) = \frac{1}{\left(\mathbb{M}^{(n)} \right)_{i,i}} \sum_{j=0}^{n-1} \left(\mathbb{J}_{\mathbb{A}}^j \mathbb{M} \bar{\mathbb{J}}_{\mathbb{A}}^j \right)_{i,i} \omega_{\zeta} \left(\bar{\mathbb{J}}_{\mathbb{A}}^j \mathbf{z}' \right). \quad (3.56)$$

É importante notar que $\bar{\mathbb{J}}_{\mathbb{A}}^j \mathbf{z}'$ continua sendo um vetor, a menos que escolhamos uma componente S_i associada a um bloco de Jordan de tamanho 1. Vemos que, por mais que a forma canônica de Jordan pareça ser bastante simples, a forma geral da solução continua sendo bastante extensa caso a matriz \mathbb{A} não seja diagonalizável. Neste caso, parece mais difícil comparar a função característica de uma única componente S_i e determinar se a distribuição limite será gaussiana, baseado apenas nos autovalores J_{q_i} , dado que os termos ainda envolvem os binômios de Newton, que parecem tornar a análise complicada. Na seção 4.3, veremos que podemos, sim, determinar a gaussianização da distribuição limite em termos do autovalor do bloco de Jordan.

4 Gaussianização com ruído independente

Agora, vamos considerar o mapa p -dimensional

$$\mathbf{x}[n+1] = \mathbb{A}\mathbf{x}[n] + \eta[n+1], \quad n \in \{0, 1, 2, \dots, N\} \quad (4.1)$$

com as mesmas suposições apresentadas no Capítulo 3, ou seja, que o ruído η é centralizado, estacionário, e independente para instantes de tempo diferentes. Porém, além disso, vamos considerar mais uma simplificação, supondo que o vetor principal de η tem suas componentes independentes entre si. Vamos supor ainda a condição inicial $\mathbf{x}[0] = 0$, apenas por simplicidade, uma vez que a condição inicial deve apenas deslocar o valor médio da distribuição, sem alterar a sua forma.

A análise a ser realizada neste capítulo é importante, como já foi discutido, porque, caso demonstramos a gaussianização de cada uma das componentes do sistema, estaremos também demonstrando a gaussianização do sistema como um todo, ou seja, que a distribuição limite do sistema é uma distribuição normal p -dimensional. Como supomos que o vetor principal do ruído tem suas componentes independentes entre si, podemos calcular a função característica em outras bases, sem nos preocuparmos com correlações que possam impedir ou dificultar a mudança de base. Além disso, ao mudarmos de base, cada componente na nova base pode ser descrita como uma combinação linear das componentes do sistema na base do vetor principal do ruído, que devem ser gaussianas, todas as componentes da nova base também deve ser gaussianas. Dado que isso deve ser válido para qualquer base, temos uma distribuição limite gaussiana em todas as direções do espaço p -dimensional, ou seja, uma distribuição normal p -dimensional.

Sendo \mathbb{T} a transformação que leva η ao seu vetor principal, que vamos chamar ζ , temos:

$$\mathbf{y}[n+1] = \mathbb{B}\mathbf{y}[n] + \zeta[n+1], \quad (4.2)$$

sendo $\mathbf{y}[n] = \mathbb{T}\mathbf{x}[n]$, $\mathbb{B} = \mathbb{T}\mathbb{A}\bar{\mathbb{T}}$, lembrando que $\mathbb{T}\bar{\mathbb{T}} = \bar{\mathbb{T}}\mathbb{T} = \mathbb{1}$. Estamos, então, supondo que as componentes de ζ devem ser independentes entre si. A matriz de covariância de ζ deve ser dada por $\mathbb{T}\mathbb{M}\bar{\mathbb{T}} = \mathbb{D}_{\mathbb{M}}^2$, $\mathbb{D}_{\mathbb{M}} = \text{diag}(M_1, M_2, \dots, M_p)$, sendo \mathbb{M} a matriz de covariância do ruído original η , e M_i^2 seus autovalores. Vamos estudar o mapa através da equação (4.2), lembrando que, eventualmente, podemos retornar à base original:

$$\psi_{\mathbf{x}[n]}(\mathbf{z}) = \psi_{\mathbf{y}[n]}(\mathbb{T}\mathbf{z}). \quad (4.3)$$

Primeiramente, descrevendo o n -ésimo passo do mapa através do esquema triangular de Kolmogorov:

$$\mathbf{y}[n] = \sum_{j=0}^{n-1} \mathbb{B}^{n-j} \zeta[j] = \sum_{j=0}^{n-1} \mathbb{B}^j \zeta[j], \quad (4.4)$$

lembrando que a segunda igualdade vem da estacionariedade do ruído. Continuando, podemos separar as equações para cada componente:

$$y_i[n] = \sum_{j=0}^{n-1} \left(\mathbb{B}^j \zeta[j] \right)_i = \sum_{j=0}^{n-1} \sum_{k=1}^p \left(\mathbb{B}^j \right)_{i,k} \zeta_k[j] = \sum_{k=1}^p S_{i,k}[n], \quad S_{i,k}[n] = \sum_{j=0}^{n-1} \left(\mathbb{B}^j \right)_{i,k} \zeta_k[j]. \quad (4.5)$$

Sendo M_k o desvio padrão da componente ζ_k , podemos calcular a variância de ambos os lados, obtendo:

$$\nu_{i,k}^2[n] = \langle (S_{i,k}[n])^2 \rangle = \sum_{j=0}^{n-1} \left(\mathbb{B}^j \right)_{i,k}^2 M_k^2 = M_k^2 \sum_{j=0}^{n-1} \left(\mathbb{B}^j \right)_{i,k}^2. \quad (4.6)$$

Dado que ζ é um vetor aleatório principal, a função característica de uma componente sua pode ser escrita:

$$\psi_{\zeta_k[n]}(z) = \psi_k(z) = \exp \left[-\frac{M_k^2 z^2}{2} (1 + \omega_k(M_k z)) \right]. \quad (4.7)$$

Com isto, escrevemos a função característica de $S_{i,k}$:

$$\begin{aligned} \psi_{i,k}(n, z) &= \langle \exp(Iz S_{i,k}[n]) \rangle = \left\langle \exp \left(Iz \sum_{j=0}^{n-1} \left(\mathbb{B}^j \right)_{i,k} \zeta_k[j] \right) \right\rangle \\ &= \prod_{j=0}^{n-1} \left\langle \exp \left(Iz \left(\mathbb{B}^j \right)_{i,k} \zeta_k[j] \right) \right\rangle \\ &= \prod_{j=0}^{n-1} \exp \left[-\frac{M_k^2 z^2}{2} \left(\mathbb{B}^j \right)_{i,k}^2 \left(1 + \omega_k \left(M_k \left(\mathbb{B}^j \right)_{i,k} z \right) \right) \right] \\ &= \exp \left\{ -\frac{M_k^2 z^2}{2} \sum_{j=0}^{n-1} \left(\mathbb{B}^j \right)_{i,k}^2 \left[1 + \omega_k \left(\left(\mathbb{B}^j \right)_{i,k} M_k z \right) \right] \right\} \\ &= \exp \left\{ -\frac{\nu_{i,k}^2[n] z^2}{2} \left[1 + \frac{M_k^2}{\nu_{i,k}^2[n]} \sum_{j=0}^{n-1} \left(\mathbb{B}^j \right)_{i,k}^2 \omega_k \left(\left(\mathbb{B}^j \right)_{i,k} M_k z \right) \right] \right\}. \end{aligned} \quad (4.8)$$

Vamos encontrar as condições para a gaussianização de $S_{i,k}$, ou seja, para que $\lim_{n \rightarrow \infty} \nu_{i,k}^{-2}[n] M_k^2 \sum_{j=0}^{n-1} (\mathbb{B}^j)_{i,k} = 0$. Vamos considerar diferentes casos separadamente, primeiramente supondo que \mathbb{B} é diagonalizável com autovalores reais, depois, com autovalores complexos e, por fim, considerando o caso em que \mathbb{B} não é diagonalizável, e avaliamos o sistema em termos de sua forma canônica de Jordan.

4.1 Matriz diagonalizável com autovalores reais

Se \mathbb{B} é diagonalizável, podemos escrever $\mathbb{B} = \mathbb{P}\mathbb{D}\mathbb{P}^{-1}$, $\mathbb{D} = \text{diag}(B_1, B_2, \dots, B_p)$, sendo B_i o i -ésimo autovalor de \mathbb{B} . Com isto, escrevemos:

$$\begin{aligned} (\mathbb{B}^j)_{i,k} &= (\mathbb{P}\mathbb{D}\mathbb{P}^{-1})_{i,k} = \sum_{m=1}^p \mathbb{P}_{i,m} (\mathbb{D}^j \mathbb{P}^{-1})_{m,k} = \sum_{m_1, m_2}^p \mathbb{P}_{i, m_1} \mathbb{D}_{m_1, m_2}^j \mathbb{P}_{m_2, k}^{-1} \\ &= \sum_{m_1, m_2}^p \mathbb{P}_{i, m_1} B_{m_1}^j \delta_{m_1, m_2} \mathbb{P}_{m_2, k}^{-1} = \sum_{m=1}^p \mathbb{P}_{i, m} \mathbb{P}_{m, k}^{-1} B_m^j. \end{aligned} \quad (4.9)$$

$$\therefore (\mathbb{B}^j)_{i,k}^2 = \sum_{m=1}^p (\mathbb{P}_{i, m} \mathbb{P}_{m, k}^{-1})^2 B_m^{2j} + 2 \sum_{m_1 < m_2} \mathbb{P}_{i, m_1} \mathbb{P}_{m_1, k} \mathbb{P}_{i, m_2} \mathbb{P}_{m_2, k}^{-1} (B_{m_1} B_{m_2})^j. \quad (4.10)$$

Vamos avaliar a gaussianização em termos dos módulos dos autovalores. Primeiro, vamos supor $|B_m| < 1 \forall m$. Neste caso, temos:

$$\sum_{j=0}^{n-1} B_m^{2j} = \frac{1 - B_m^{2n}}{1 - B_m^2} \xrightarrow{n \rightarrow \infty} \frac{1}{1 - B_m^2}; \quad (4.11)$$

$$\sum_{j=0}^{n-1} (B_{m_1} B_{m_2})^j = \frac{1 - (B_{m_1} B_{m_2})^n}{1 - (B_{m_1} B_{m_2})} \xrightarrow{n \rightarrow \infty} \frac{1}{1 - B_{m_1} B_{m_2}}. \quad (4.12)$$

Portanto:

$$\begin{aligned} M_k^{-2} \nu_{i,k}^2[n] &= \sum_{m=1}^p (\mathbb{P}_{i, m} \mathbb{P}_{m, k}^{-1})^2 \frac{1 - B_m^{2n}}{1 - B_m^2} + 2 \sum_{m_1 < m_2} \mathbb{P}_{i, m_1} \mathbb{P}_{m_1, k}^{-1} \mathbb{P}_{i, m_2} \mathbb{P}_{m_2, k}^{-1} \frac{1 - (B_{m_1} B_{m_2})^n}{1 - (B_{m_1} B_{m_2})} \\ &\xrightarrow{n \rightarrow \infty} \sum_{m=1}^p \frac{(\mathbb{P}_{i, m} \mathbb{P}_{m, k}^{-1})^2}{1 - B_m^2} + 2 \sum_{m_1 < m_2} \frac{\mathbb{P}_{i, m_1} \mathbb{P}_{m_1, k}^{-1} \mathbb{P}_{i, m_2} \mathbb{P}_{m_2, k}^{-1}}{1 - (B_{m_1} B_{m_2})} < \infty. \end{aligned} \quad (4.13)$$

$$\therefore \lim_{n \rightarrow \infty} \frac{M_k^2 \max_{0 \leq j \leq n-1} ((\mathbb{B}^j)_{i,k}^2)}{\nu_{i,k}^2[n]} = \frac{M_k^2 \max_{0 \leq j \leq n-1} ((\mathbb{B}^j)_{i,k}^2)}{\nu_{i,k}^2} > 0, \quad (4.14)$$

a não ser que $(\mathbb{B}^j)_{i,k}^2 = 0 \forall j \in \mathcal{N}$. Apesar de parecer improvável, veremos que pode acontecer. Note que definimos $\nu_{i,k}^2 = \lim_{n \rightarrow \infty} \nu_{i,k}^2[n]$. Portanto, concluímos que, se os módulos de todos os autovalores são menores que 1, a distribuição limite não é gaussiana, uma vez que sua função característica terá contribuições da função ω_k . É claro que este resultado depende de os termos $\mathbb{P}_{i,m} \mathbb{P}_{m,k}$ anularem ou não a contribuição dos autovalores. Se as contribuições de todos os autovalores fossem anuladas, os termos com ω_k seriam anulados, e a distribuição do sistema seria gaussiana.

Agora, vamos considerar o caso $|B_m| > 1 \forall m$. Novamente, podemos escrever fórmulas para as somas das séries geométricas, porém, desta vez elas divergem:

$$\sum_{j=0}^{n-1} B_m^{2j} = \frac{1 - B_m^{2n}}{1 - B_m^2} \xrightarrow{n \rightarrow \infty} \infty; \quad (4.15)$$

$$\sum_{j=0}^{n-1} (B_{m_1} B_{m_2})^j = \frac{1 - (B_{m_1} B_{m_2})^n}{1 - (B_{m_1} B_{m_2})} \xrightarrow{n \rightarrow \infty} \infty. \quad (4.16)$$

Neste caso, é claro que $\max_{0 \leq j \leq n-1} ((\mathbb{B}^j)_{i,k}^2) = (\mathbb{B}^{n-1})_{i,k}^2$ se $n \gg 1$. Além disso, podemos escrever:

$$(\mathbb{B}^{n-1})_{i,k}^2 \xrightarrow{n \rightarrow \infty} \mathbb{P}_{i,m_*} \mathbb{P}_{m_*,k}^{-1} B_{m_*}^{2(n-1)}, \quad (4.17)$$

sendo m_* tal que B_{m_*} é o autovalor com o maior módulo entre aqueles cujas contribuições não são anuladas pelos termos $\mathbb{P}_{i,m} \mathbb{P}_{m,k}^{-1}$ que os acompanham. Note que também podemos escrever:

$$M_k^{-2} \nu_{i,k}^2[n] \xrightarrow{n \rightarrow \infty} \mathbb{P}_{i,m_*} \mathbb{P}_{m_*,k}^{-1} \frac{B_{m_*}^{2n}}{B_{m_*}^2 - 1}. \quad (4.18)$$

Deste modo, obtemos a razão:

$$\lim_{n \rightarrow \infty} \frac{M_k^2 (\mathbb{B}^j)_{i,k}^2}{\nu_{i,k}^2[n]} = \frac{B_{m_*}^2 - 1}{B_{m_*}^2} > 0. \quad (4.19)$$

Concluímos que, quando $|B_m| > 1 \forall m$, a distribuição limite do sistema também não é gaussiana, dado que a contribuição dos autovalores não sejam todas anuladas. Note que, além do caso $\mathbb{P}_{i,m} \mathbb{P}_{m,k}^{-1} = 0$, há outra possibilidade de anulação de um autovalor: caso existam m_1 e m_2 diferentes entre si tais que $B_{m_1} = B_{m_2}$ e

$$\mathbb{P}_{i,m_1} \mathbb{P}_{m_1,k}^{-1} = -\mathbb{P}_{i,m_2} \mathbb{P}_{m_2,k}^{-1}, \quad (4.20)$$

então as contribuições de B_{m_1} e B_{m_2} se anulam conjuntamente.

Finalmente, vamos avaliar o caso $|B_m| = 1 \forall m$. Escrevendo:

$$(\mathbb{B}^j)_{i,k}^2 \leq \sum_{m=1}^p (\mathbb{P}_{i,m} \mathbb{P}_{m,k}^{-1})^2 + 2 \sum_{m_1 \leq m_2} |\mathbb{P}_{i,m_1} \mathbb{P}_{m_1,k}^{-1} \mathbb{P}_{i,m_2} \mathbb{P}_{m_2,k}^{-1}|; \quad (4.21)$$

$$\begin{aligned} M_k^{-2} \nu_{i,k}^2[n] &= \sum_{j=0}^{n-1} (\mathbb{B}^j)_{i,k}^2 \\ &= n \sum_{m=1}^p (\mathbb{P}_{i,m} \mathbb{P}_{m,k}^{-1})^2 + 2 \sum_{m_1 \leq m_2} \mathbb{P}_{i,m_1} \mathbb{P}_{m_1,k}^{-1} \mathbb{P}_{i,m_2} \mathbb{P}_{m_2,k}^{-1} \sum_{j=0}^{n-1} (B_{m_1} B_{m_2})^j \\ &\geq n \sum_{m=1}^p (\mathbb{P}_{i,m} \mathbb{P}_{m,k}^{-1})^2 + 2n \sum_{B_{m_1} = B_{m_2}} \mathbb{P}_{i,m_1} \mathbb{P}_{m_1,k}^{-1} \mathbb{P}_{i,m_2} \mathbb{P}_{m_2,k}^{-1}, \end{aligned} \quad (4.22)$$

onde a última desigualdade vem dos termos envolvendo m_1 e m_2 tais que $B_{m_1} = -B_{m_2}$, de modo que

$$\sum_{j=0}^{n-1} (B_{m_1} B_{m_2})^j = 0 \text{ ou } 1.$$

Com isto, é claro que

$$\lim_{n \rightarrow \infty} \frac{M_k^2 (\mathbb{B}^j)_{i,k}^2}{\nu_{i,k}^2} = 0. \quad (4.23)$$

Portanto, se os autovalores têm módulo unitário, a distribuição limite será gaussiana. Note que, caso os autovalores sejam anulados pelos termos das matrizes de mudança de base, suas contribuições para a função característica seriam também anuladas junto com as funções ω , mantendo a gaussianidade da distribuição limite. Note que podemos apresentar resultados um pouco mais gerais, considerando apenas o autovalor de maior módulo cuja contribuição não é anulada.

Considerando apenas os autovalores cujas contribuições não são anuladas, se o maior módulo é maior que 1, novamente teremos o resultado da Equação (4.19), uma vez que os efeitos dos autovalores de menor módulo se tornam desprezíveis. Do mesmo modo, se o maior módulo for igual a 1, o numerador $M_k^2 (\mathbb{B}^j)_{i,k}^2$ será finito, enquanto o denominador $\nu_{i,k}^2[n]$ será proporcional a n , com as contribuições dos autovalores menores sendo limitadas, ou seja, comparativamente desprezíveis. Então, o denominador tende a infinito com n e, portanto, a razão tende a zero, resultando na gaussianização da distribuição.

Vamos resumir as conclusões obtidas para o caso da matriz \mathbb{B} diagonalizável com autovalores reais. Consideremos os autovalores B_m de \mathbb{B} , sendo m tal que $\mathbb{P}_{i,m}\mathbb{P}_{m,k}^{-1} \neq 0$. Caso existam $m_1 \neq m_2$ tais que $B_{m_1} = B_{m_2}$, se $\mathbb{P}_{i,m_1}\mathbb{P}_{m_1,k}^{-1} \neq -\mathbb{P}_{i,m_2}\mathbb{P}_{m_2,k}^{-1}$, então incluímos m_1 e m_2 nos possíveis valores de m . Temos, então:

1. Se $0 < \max_{1 \leq m \leq p} (|B_m|) = 1$, a distribuição limite de $S_{i,k}$ é gaussiana;
2. Caso contrário, isto é, $\max_{1 \leq m \leq p} (|B_m|) < 1$ ou $\max_{1 \leq m \leq p} (|B_m|) > 1$, a distribuição limite **não** é gaussiana.

4.2 Matriz diagonalizável com autovalores complexos

Aqui, vamos fazer uma análise muito semelhante àquela apresentada na seção anterior. É importante notar que as séries geométricas não devem convergir para limites bem definidos quando consideramos autovalores complexos de módulo maior ou igual a 1. Para contornar isso, utilizaremos o fato de que, se um número complexo é autovalor da matriz, então seu complexo conjugado também deve ser. Por simplicidade, na continuação do texto, vamos considerar que, quando fazemos suposições sobre os autovalores, estamos considerando apenas aqueles cujas contribuições não são anuladas pelas matrizes de mudança de base.

Primeiramente, se $|B_m| < 1 \forall m$, as Equações (4.11) e (4.12) se repetem, e novamente temos o limite

$$\lim_{n \rightarrow \infty} \frac{M_k^2 \max \left((\mathbb{B}^j)_{i,k}^2 \right)}{\nu_{i,k}^2[n]} > 0, \quad (4.24)$$

de modo que a distribuição limite não é gaussiana também para autovalores complexos, caso todos tenham módulo menor que 1.

Para $|B_m| > 1 \forall m$, sendo m_1 e m_2 são tais que B_{m_1} e B_{m_2} são complexos conjugados e ambos possuem o maior módulo entre os autovalores do sistema. Com isto, podemos escrever

$$\left(\mathbb{B}^{n-1} \right)_{i,k}^2 \xrightarrow{n \rightarrow \infty} \left(\mathbb{P}_{i,m_1} \mathbb{P}_{m_1,k}^{-1} B_{m_1}^{n-1} + \mathbb{P}_{i,m_2} \mathbb{P}_{m_2,k}^{-1} B_{m_2}^{n-1} \right)^2, \quad (4.25)$$

a menos que o lado direito seja nulo, o que só é possível caso $\left| \mathbb{P}_{i,m_1} \mathbb{P}_{m_1,k}^{-1} \right| = \left| \mathbb{P}_{i,m_2} \mathbb{P}_{m_2,k}^{-1} \right|$ e, sendo $B_{m_1} = r e^{I\theta}$, caso exista n tal que:

$$\begin{aligned} \operatorname{Re} \left(\mathbb{P}_{i,m_1} \mathbb{P}_{m_1,k}^{-1} + \mathbb{P}_{i,m_2} \mathbb{P}_{m_2,k}^{-1} \right) \cos((n-1)\theta) &= \operatorname{Im} \left(\mathbb{P}_{i,m_1} \mathbb{P}_{m_1,k}^{-1} - \mathbb{P}_{i,m_2} \mathbb{P}_{m_2,k}^{-1} \right) \sin((n-1)\theta); \\ \operatorname{Im} \left(\mathbb{P}_{i,m_1} \mathbb{P}_{m_1,k}^{-1} + \mathbb{P}_{i,m_2} \mathbb{P}_{m_2,k}^{-1} \right) \cos((n-1)\theta) &= \operatorname{Re} \left(\mathbb{P}_{i,m_2} \mathbb{P}_{m_2,k}^{-1} - \mathbb{P}_{i,m_1} \mathbb{P}_{m_1,k}^{-1} \right) \sin((n-1)\theta). \end{aligned}$$

Isto implica que a distribuição limite de $S_{i,k}[n]$ pode apresentar um caráter oscilatório. Porém, isso não necessariamente será perceptível na distribuição de x_i , uma vez que ela também deve depender de termos $S_{i,k'}$, $k' \neq k$.

De qualquer modo, ainda podem haver autovalores que não sejam anulados para o mesmo n , de modo que $(\mathbb{B}^{n-1})_{i,k}^2$ seja proporcional a $r_\star = |B_{m_\star}| = |B_{m'_\star}|$, sendo B_{m_\star} o autovalor de maior módulo, junto de seu complexo conjugado $B_{m'_\star}$, tais que suas contribuições não são anuladas para $j = n - 1$. Explicitamente:

$$\left(\mathbb{B}^{n-1} \right)_{i,k}^2 \xrightarrow{n \rightarrow \infty} \left(\mathbb{P}_{i,m_\star} \mathbb{P}_{m_\star,k} e^{In\theta_\star} + \mathbb{P}_{i,m'_\star} \mathbb{P}_{m'_\star,k} e^{-In\theta_\star} \right)^2 r_\star^{2(n-1)}. \quad (4.26)$$

Da mesma maneira, para a variância da soma, escrevemos:

$$M_k^{-2} \nu_{i,k}^2[n] \xrightarrow{n \rightarrow \infty} \left(\mathbb{P}_{i,m_\star} \mathbb{P}_{m_\star,k} e^{In\theta_\star} + \mathbb{P}_{i,m'_\star} \mathbb{P}_{m'_\star,k} e^{-In\theta_\star} \right)^2 \frac{r_\star^{2n} - 1}{r_\star^2 - 1}. \quad (4.27)$$

Concluimos que a razão das variâncias no limite é dada por

$$\lim_{n \rightarrow \infty} \frac{M_k^2 (\mathbb{B}^j)_{i,k}^2}{\nu_{i,k}^2[n]} = \frac{r_\star^2 - 1}{r_\star^2} > 0, \quad (4.28)$$

ou seja, para autovalores complexos com módulo maior que 1 a distribuição limite de $S_{i,k}$ não é gaussiana, dadas as restrições associadas às matrizes de mudança de base.

Por fim, para autovalores de módulo unitário, $|B_m| = 1 \forall m$, sabemos que as potências de cada um devem apresentar caráter oscilatório. Porém, considerando pares de autovalores complexos conjugados B_{m_1} e B_{m_2} , sabemos que a soma das potências do produto entre eles deve ser

$$\sum_{j=0}^{n-1} (B_{m_1} B_{m_2})^j = n. \quad (4.29)$$

Assim, sabemos que $\nu_{i,k}^2[n]$, apesar de apresentar termos oscilatórios, estes devem ter amplitude limitada, apresenta também termos proporcionais a n de modo que, para $n \gg 1$,

$\nu_{i,k}^2[n]$ deve ser proporcional a n , enquanto $(\mathbb{B}^j)_{i,k}^2$ tem seu módulo limitado. Note que, pela desigualdade triangular, podemos afirmar que o módulo de $(\mathbb{B}^j)_{i,k}^2$ deve ser menor ou igual à soma dos módulos dos termos apresentados na Equação (4.10), o que claramente é limitado. Com isso, concluímos que

$$\lim_{n \rightarrow \infty} \frac{M_k^2 (\mathbb{B}^j)_{i,k}^2}{\nu_{i,k}^2[n]} = 0, \quad (4.30)$$

para $|B_m| = 1 \forall m$, ou seja, a distribuição limite de $S_{i,k}[n]$ será gaussiana caso os autovalores tenham módulo unitário, mesmo que sejam complexos. Note que, assim como no caso dos autovalores reais, será determinante na gaussianidade da distribuição limite apenas o maior módulo entre os autovalores com contribuições não eliminadas quando consideramos os termos de \mathbb{P} e \mathbb{P}^{-1} .

4.3 Matriz não Diagonalizável: Forma Canônica de Jordan

Agora, vamos considerar o caso em que não podemos escrever a matriz \mathbb{B} em termos de uma matriz diagonal. Na verdade, o melhor que podemos fazer é escrever \mathbb{B} em função de sua forma canônica de Jordan $\mathbb{J}_{\mathbb{B}}$:

$$\mathbb{B} = \mathbb{P} \mathbb{J}_{\mathbb{B}} \mathbb{P}^{-1}. \quad (4.31)$$

Note que podemos escrever a matriz de Jordan como uma matriz diagonal, de modo que cada uma das entradas da diagonal são blocos de Jordan: $\mathbb{J}_{\mathbb{B}} = \text{diag}(\mathbb{J}_{(1)}, \mathbb{J}_{(2)}, \dots, \mathbb{J}_{(r)})$, ou seja, a matriz tem r blocos de Jordan, cada um de tamanho p_i , de modo que $p_1 + p_2 + \dots + p_r = p$. Desta maneira, podemos escrever $\mathbb{J}_{\mathbb{B}}^n = \text{diag}(\mathbb{J}_{(1)}^n, \mathbb{J}_{(2)}^n, \dots, \mathbb{J}_{(r)}^n)$, e já sabemos como calcular a potência de um bloco de Jordan. Com isto, sabemos qual deve ser o valor de cada célula das potências de \mathbb{J} :

$$(\mathbb{J}^n)_{i,j} = \begin{cases} (\mathbb{J}_{(s)}^n)_{i-p'_{s-1}, j-p'_{s-1}}, & p'_{s-1} < i, j \leq p'_s; \\ 0, & \text{caso contrário.} \end{cases} \quad (4.32)$$

Ou seja, temos os termos das potências dos blocos de Jordan, e o resto da matriz é nulo. Aqui, definimos:

$$p'_s = \sum_{i=0}^s p_i, \quad p_0 = 0. \quad (4.33)$$

Com isto, calculamos:

$$\begin{aligned}
(\mathbb{B}^j)_{i,k} &= (\mathbb{P} \mathbb{J}_{\mathbb{B}} \mathbb{P}^{-1})_{i,k} \sum_{m_1, m_2=1}^p \mathbb{P}_{i, m_1} (\mathbb{J}_{\mathbb{B}}^j)_{m_1, m_2} \mathbb{P}_{m_2, k}^{-1} \\
&= \sum_{q=1}^r \left\{ \sum_{m_1, m_2=p'_q-1}^{p'_q} \mathbb{P}_{i, m_1} \mathbb{P}_{m_2, k}^{-1} (\mathbb{J}_{(q)}^j)_{m_1-p'_{q-1}, m_2-p'_{q-1}} \right\} \\
&= \sum_{q=1}^r \left\{ \sum_{m_1, m_2=1}^{p_q} \mathbb{P}_{i, m_1+p'_{q-1}} \mathbb{P}_{m_2+p'_{q-1}, k}^{-1} (\mathbb{J}_{(q)}^j)_{m_1, m_2} \right\}. \tag{4.34}
\end{aligned}$$

Lembrando que os termos da potência de um bloco de Jordan são dados por

$$(\mathbb{J}_{(q)}^j)_{i,k} = \sum_{a=0}^{W_q} \binom{j}{a} J_q^{j-a} \delta_{i, k-a}, \quad W_q = \min \{p_q - 1, j\} \tag{4.35}$$

sendo J_q o autovalor do bloco, continuamos:

$$\begin{aligned}
(\mathbb{B}^j)_{i,k} &= \sum_{q=1}^r \left\{ \sum_{m_1, m_2=1}^{p_q} \mathbb{P}_{i, m_1+p'_{q-1}} \mathbb{P}_{m_2+p'_{q-1}, k}^{-1} \sum_{a=0}^j \binom{j}{a} J_q^{j-a} \delta_{m_1, m_2-a} \right\} \\
&= \sum_{q=1}^r \sum_{m_1, m_2=1}^{p_q} \sum_{a=0}^{W_q} \mathbb{P}_{i, m_1+p'_{q-1}} \mathbb{P}_{m_2+p'_{q-1}, k}^{-1} \binom{j}{a} J_q^{j-a} \delta_{m_1, m_2-a} \\
&= \sum_{q=1}^r \sum_{m=1}^{p_q} \sum_{a=0}^{W_q} \mathbb{P}_{i, m+p'_{q-1}} \mathbb{P}_{m+p'_{q-1}+a, k}^{-1} \binom{j}{a} J_q^{j-a}. \tag{4.36}
\end{aligned}$$

Com isto, calculamos o quadrado:

$$\begin{aligned}
(\mathbb{B}^j)_{i,k}^2 &= \sum_{q=1}^r \sum_{m=1}^{p_q} \sum_{a=0}^{W_q} \left(\mathbb{P}_{i, m+p'_{q-1}} \mathbb{P}_{m+p'_{q-1}+a, k}^{-1} \right)^2 \binom{j}{a}^2 J_q^{2(j-a)} \\
&+ 2 \sum_{\substack{q_1 < q_2 \\ \text{ou } m_1 < m_2 \\ \text{ou } a_1 < a_2}} \mathbb{P}_{i, m_1+p'_{q_1-1}} \mathbb{P}_{m_1+p'_{q_1-1}+a_1, k}^{-1} \mathbb{P}_{i, m_2+p'_{q_2-1}} \mathbb{P}_{m_2+p'_{q_2-1}+a_2, k}^{-1} \binom{j}{a_1} \binom{j}{a_2} (J_{q_1} J_{q_2})^{j-a}. \tag{4.37}
\end{aligned}$$

Note que, no último somatório, as condições apresentadas não precisam todas acontecer ao mesmo tempo, sendo apenas uma delas necessária para obter termos não considerados no primeiro somatório. Agora podemos avaliar a razão entre $M_k^2(\mathbb{B}^j)_{i,k}$ e $\nu_{i,k}^2[n]$ para diferentes conjuntos de autovalores J_q . Primeiramente, consideramos $|J_q| < 1 \forall q$. Para calcular $\nu_{i,k}^2[n]$, note que podemos escrever:

$$\sum_{j=0}^{n-1} \left(\frac{j!}{(j-a)!} \right)^2 J_q^{2(j-a)} < \sum_{j=0}^{n-1} j^{2a} J_q^{2(j-a)} = \sum_{j=0}^{n-1} \frac{J_q^{-2a}}{\ln(J_q)^{2a}} \frac{d^{2a}}{d\alpha^{2a}} (J_q^{\alpha j}) \Big|_{\alpha=2}. \tag{4.38}$$

$$\begin{aligned}
\therefore \sum_{j=0}^{n-1} \left(\frac{j!}{(j-a)!} \right)^2 J_q^{2(j-a)} &< \frac{J_q^{-2a}}{\ln(J_q)^{2a}} \frac{d^{2a}}{d\alpha^{2a}} \left(\sum_{j=0}^{n-1} J_q^{\alpha j} \right) \Bigg|_{\alpha=2} \\
\sum_{j=0}^{n-1} \left(\frac{j!}{(j-a)!} \right)^2 J_q^{2(j-a)} &< \frac{J_q^{-2a}}{\ln(J_q)^{2a}} f_n(J_q, 2a), \tag{4.39} \\
f_n(J_q, 2a) &= \frac{d^{2a}}{d\alpha^{2a}} \left(\frac{1 - J_q^{\alpha n}}{1 - J_q^\alpha} \right) \Bigg|_{\alpha=2}.
\end{aligned}$$

Podemos tratar da mesma maneira os termos do segundo somatório, que podem envolver autovalores diferentes:

$$\begin{aligned}
\binom{j}{a_1} \binom{j}{a_2} J_{q_1}^{j-a_1} J_{q_2}^{j-a_2} &< j^{a_1+a_2} J_{q_1}^{j-a_1} J_{q_2}^{j-a_2} = J_{q_1}^{-a_1} J_{q_2}^{-a_2} j^{a_1+a_2} (J_{q_1} J_{q_2})^j \\
&= \frac{J_{q_1}^{-a_1} J_{q_2}^{-a_2}}{\ln(J_{q_1} J_{q_2})^{a_1+a_2}} \frac{d^{a_1+a_2}}{d\alpha^{a_1+a_2}} \left[(J_{q_1} J_{q_2})^{\alpha j} \right] \Bigg|_{\alpha=1}. \tag{4.40}
\end{aligned}$$

Deste modo, a contribuição desses termos para $\nu_{i,k}^2[n]$ é

$$\sum_{j=0}^{n-1} \binom{j}{a_1} \binom{j}{a_2} J_{q_1}^{j-a_1} J_{q_2}^{j-a_2} < \frac{J_{q_1}^{-a_1} J_{q_2}^{-a_2}}{\ln(J_{q_1} J_{q_2})^{a_1+a_2}} g_n(J_{q_1}, J_{q_2}, a_1, a_2), \tag{4.41}$$

$$g_n(J_{q_1}, J_{q_2}, a_1, a_2) = \frac{d^{a_1+a_2}}{d\alpha^{a_1+a_2}} \left[\frac{1 - (J_{q_1} J_{q_2})^{\alpha n}}{1 - (J_{q_1} J_{q_2})^\alpha} \right] \Bigg|_{\alpha=1}. \tag{4.42}$$

Note que $f_n(J_q, 2a) = g_n(J_q, J_q, a, a)$. Note ainda que podemos escrever, de modo geral, conhecendo a regra de derivação de uma divisão:

$$g_n(J_{q_1}, J_{q_2}, a_1, a_2) = \frac{g'_n(J_{q_1}, J_{q_2}, a_1, a_2)}{(1 - J_{q_1} J_{q_2})^{2(a_1+a_2)}}, \tag{4.43}$$

onde o denominador nunca se anulará, dado que os autovalores têm módulo menor que 1, enquanto o numerador será dado por um polinômio de $(J_{q_1} J_{q_2})$ e $\ln(J_{q_1} J_{q_2})$. Tomando o limite $n \rightarrow \infty$, os termos com expoente n são anulados, e g'_n resulta num polinômio de grau máximo $2^{a_1+a_2}$. Deste modo,

$$\begin{aligned}
\nu_{i,k}^2[n] &= \sum_{q=1}^r \sum_{m=1}^{p_q} \sum_{a=0}^{W_q} \left(\mathbb{P}_{i,m+p'_{q-1}} \mathbb{P}_{m+p'_{q-1}+a,k}^{-1} \right)^2 \frac{J_q^{-2a}}{\ln(J_q)^{2a}} f_n(J_q, 2a) \tag{4.44} \\
&+ 2 \sum_{\substack{q_1 < q_2 \\ \text{ou } m_1 < m_2 \\ \text{ou } a_1 < a_2}} \mathbb{P}_{i,m_1+p'_{q_1-1}} \mathbb{P}_{m_1+p'_{q_1-1}+a_1,k}^{-1} \mathbb{P}_{i,m_2+p'_{q_2-1}} \mathbb{P}_{m_2+p'_{q_2-1}+a_2,k}^{-1} \frac{J_{q_1}^{-a_1} J_{q_2}^{-a_2} g_n(J_{q_1}, J_{q_2}, a_1, a_2)}{\ln(J_{q_1} J_{q_2})^{a_1+a_2}},
\end{aligned}$$

que deve ser limitado, mesmo no limite $n \rightarrow \infty$, se lembramos que estamos considerando o módulo dos autovalores menor que 1, mesmo que estes sejam complexos. Assim, temos o limite:

$$\lim_{n \rightarrow \infty} \frac{\max \left((\mathbb{B}^j)_{i,k} \right)}{\nu_{i,k}^2[n]} > 0, \quad (4.45)$$

e a distribuição limite não é gaussiana. É claro que ainda seria necessário considerar em detalhe os termos de \mathbb{P} e \mathbb{P}^{-1} e, caso $(\mathbb{B}^j)_{i,k} = 0 \forall j$, então a distribuição limite de $S_{i,k}$ deverá, sim, ser gaussiana.

Agora, vamos avaliar o caso $|J_q| > 1 \forall q$. A análise é muito semelhante ao caso anterior, porém, desta vez, f_n e g_n não são limitadas para $n \rightarrow \infty$, e devemos observar como $(\mathbb{B}^j)_{i,k}$ e $\nu_{i,k}^2[n]$ se balanceiam. Neste caso, para $n \gg 1$, o comportamento de $\nu_{i,k}^2[n]$ se reduzirá aos termos envolvendo o autovalor de maior módulo, J_{q^*} (ou o par, no caso de autovalores complexos), e podemos escrever:

$$\nu_{i,k}^2[n] < \sum_{m=1}^{p_{q^*}} \sum_{a=0}^{W_{q^*}} \left(\mathbb{P}_{i,m+p'_{q^*-1}} \mathbb{P}_{m+p'_{q^*-1}+a,k}^{-1} \right)^2 \frac{J_{q^*}^{-2a}}{\ln(J_{q^*})^{2a}} f_n(J_{q^*}, 2a). \quad (4.46)$$

Note que, em f'_n , deverá haver um termo que multiplica n^{2a} , dada a regra de derivação do quociente. Explicitamente, esse termo é dado por:

$$n^{2a} J_{q^*}^{2n} \prod_{l=0}^{2a-1} (J_{q^*}^2 - 1)^{2l} = n^{2a} J_{q^*}^{2n} (J_{q^*}^2 - 1)^{2^{2a}-1}. \quad (4.47)$$

Qualquer outro termo de f'_n deve apresentar uma potência menor de n , mesmo que ainda possa envolver $J_{q^*}^{2n}$, tornando-se comparativamente desprezível. Portanto, podemos dizer que, no limite $n \rightarrow \infty$, $f'_n(J_{q^*}, 2a)$ é descrito pela Equação (4.47). Esta consideração será importante para analisarmos a gaussianidade da distribuição limite. Para $n \gg 1$, $(\mathbb{B}^{n-1})_{i,k}^2$ também se reduz aos termos envolvendo o autovalor de maior módulo:

$$\begin{aligned} (\mathbb{B}^{n-1})_{i,k}^2 &= \sum_{m=1}^{p_{q^*}} \sum_{a=0}^{W_{q^*}} \left(\mathbb{P}_{i,m+p'_{q^*-1}} \mathbb{P}_{m+p'_{q^*-1}+a,k}^{-1} \right)^2 \binom{n-1}{a} J_{q^*}^{2(n-1-a)} \\ &> \sum_{m=1}^{p_{q^*}} \sum_{a=0}^{W_{q^*}} \left(\mathbb{P}_{i,m+p'_{q^*-1}} \mathbb{P}_{m+p'_{q^*-1}+a,k}^{-1} \right)^2 \frac{(n-1-a)^{2a}}{(a!)^2} J_{q^*}^{2(n-1-a)}. \end{aligned} \quad (4.48)$$

Vamos avaliar a razão entre termos de $(\mathbb{B}^{n-1})_{i,k}^2$ e $\nu_{i,k}^2[n]$ para um mesmo valor de a :

$$\begin{aligned}
\lim_{n \rightarrow \infty} \frac{(n-1-a)^{2a} J_{q_*}^{2(n-1-a)}}{J_{q_*}^{-2a} f_n(J_{q_*}, 2a)} &= \lim_{n \rightarrow \infty} \frac{n^{2a} J_{q_*}^{2(n-1)}}{f_n(J_{q_*}, 2a)} (J_{q_*}^2 - 1)^{2a} \\
&= \lim_{n \rightarrow \infty} n^{2a} J_{q_*}^{2(n-1)} \frac{(J_{q_*}^2 - 1)^{2a}}{n^{2a} J_{q_*}^{2n} (J_{q_*}^2 - 1)^{2a-1}} \\
&= \frac{J_{q_*}^2 - 1}{J_{q_*}^2} > 0.
\end{aligned} \tag{4.49}$$

Perceba que o resultado não depende do valor de a . Agora, note que, para dois conjuntos de números $a_i \geq 0$ e $b_i > 0 \forall i = 1, 2, \dots, n$, podemos escrever:

$$\frac{a_i}{b_i} \geq \min_{1 \leq i \leq n} \left(\frac{a_i}{b_i} \right) \Rightarrow a_i \geq b_i \min_{1 \leq i \leq n} \left(\frac{a_i}{b_i} \right) \Rightarrow \sum_{i=1}^n a_i \geq \sum_{i=1}^n b_i \min_{1 \leq i \leq n} \left(\frac{a_i}{b_i} \right). \tag{4.50}$$

$$\therefore \frac{\sum_{i=1}^n a_i}{\sum_{i=1}^n b_i} \geq \min_{1 \leq i \leq n} \left(\frac{a_i}{b_i} \right). \tag{4.51}$$

Considerando que o autovalores não têm suas contribuições anuladas por termos de \mathbb{P} e \mathbb{P}^{-1} , temos:

$$\begin{aligned}
&\lim_{n \rightarrow \infty} \min_{\substack{1 \leq m \leq p_{q_*} \\ 0 \leq a < p_{q_*}}} \left(\frac{\left(\mathbb{P}_{i, m+p'_{q_*-1}} \mathbb{P}_{m+p'_{q_*-1}+a, k}^{-1} \right)^2 \frac{(n-1-a)^{2a} J_{q_*}^{2(n-1-a)}}{(a!)^2}}{\left(\mathbb{P}_{i, m+p'_{q_*-1}} \mathbb{P}_{m+p'_{q_*-1}+a, k}^{-1} \right)^2 \frac{J_{q_*}^{-2a}}{\ln(J_{q_*})^{2a}} f_n(J_{q_*}, 2a)} \right) \\
&= \lim_{n \rightarrow \infty} \min_{0 \leq a < p_{q_*}} \left(\frac{\frac{(n-1-a)^{2a} J_{q_*}^{2(n-1-a)}}{(a!)^2}}{\frac{J_{q_*}^{-2a}}{\ln(J_{q_*})^{2a}} f_n(J_{q_*}, 2a)} \right) \\
&= \min_{0 \leq a < p_{q_*}} \left(\frac{\ln(J_{q_*})^{2a} J_{q_*}^2 - 1}{(a!)^2 J_{q_*}} \right) = \frac{J_{q_*}^2 - 1}{J_{q_*}} \min_{0 \leq a < p_{q_*}} \left(\frac{\ln(J_{q_*})^{2a}}{(a!)^2} \right).
\end{aligned} \tag{4.52}$$

Portanto, concluímos que

$$\lim_{n \rightarrow \infty} \frac{(\mathbb{B}^{n-1})_{i, k}^2}{\nu_{i, k}^2[n]} > \frac{J_{q_*}^2 - 1}{J_{q_*}} \min_{0 \leq a < p_{q_*}} \left(\frac{\ln(J_{q_*})^{2a}}{(a!)^2} \right) > 0, \tag{4.53}$$

e a distribuição limite não será gaussiana, dadas condições sobre \mathbb{P} para que a contribuição dos autovalores não sejam anuladas.

Por fim, consideramos o caso $|J_q| = 1 \forall q$. Neste caso, para $(\mathbb{B}^j)_{i,k}^2$, podemos escrever:

$$\binom{j}{a}^2 < \frac{j^{2a}}{(a!)^2}, \quad (4.54)$$

que deve ser máximo para $j = n - 1$, é claro. Isto será utilizado para determinarmos o limite infinito da razão entre $(\mathbb{B}^j)_{i,k}^2$ e $\nu_{i,k}^2[n]$. Caso o autovalor seja complexo, ao invés de olharmos para os termos envolvendo potências de um mesmo autovalor, podemos considerar os pares de autovalores conjugados, cujas potências não devem oscilar, e os resultados serão os mesmos. Precisamos calcular o somatório em j de $\nu_{i,k}^2[n]$. Considerando um autovalor real (o que é o mesmo que considerar um par de autovalores complexos conjugados), veja:

$$\begin{aligned} \sum_{j=0}^{n-1} \binom{j}{a}^2 &> \frac{1}{(a!)^2} \sum_{j=0}^{n-1} (j-a)^{2a} = \frac{1}{(a!)^2} \sum_{j=0}^{n-1} \sum_{l=0}^{2a} \binom{2a}{l} j^{2a-l} a^l \\ &= \frac{1}{(a!)^2} \sum_{l=0}^{2a} \binom{2a}{l} a^l \sum_{j=0}^{n-1} j^{2a-l}. \end{aligned} \quad (4.55)$$

Agora, note que

$$\sum_{j=0}^{n-1} j^{2a-l} > \int_0^{n-1} x^{2a-l} dx = \frac{(n-1)^{2a-l+1}}{2a-l+1}. \quad (4.56)$$

Portanto, temos:

$$\begin{aligned} \sum_{j=0}^{n-1} \binom{j}{a}^2 &> \frac{1}{(a!)^2} \sum_{l=0}^{2a} \binom{2a}{l} a^l \frac{(n-1)^{2a-l+1}}{2a-l+1} \\ &> \frac{n-1}{(a!)^2} \sum_{l=0}^{2a} \binom{2a}{l} a^l \frac{(n-1)^{2a-l+1}}{2a-l+1} \end{aligned} \quad (4.57)$$

que podemos ainda continuar de modo a simplificar a expressão:

$$\begin{aligned} \sum_{j=0}^{n-1} \binom{j}{a}^2 &> \frac{n-1}{(a!)^2} \sum_{l=0}^{2a} \binom{2a}{l} a^l \frac{(n-1)^{2a-l}}{2a+1} = \frac{(n-1)(n-1+a)^{2a}}{(2a+1)(a!)^2} \\ &> \frac{(n-1)^{2a+1}}{(2a+1)(a!)^2}. \end{aligned} \quad (4.58)$$

Então, avaliamos a razão entre os termos de $(\mathbb{B}^{n-1})_{i,k}^2$ e $\nu_{i,k}^2[n]$ para diferentes valores de a :

$$\frac{\binom{n-1}{a}^2}{\sum_{j=0}^{n-1} \binom{j}{a}^2} < \frac{(n-1)^{2a}}{(a!)^2} \left[\frac{(n-1)^{2a+1}}{(2a+1)(a!)^2} \right]^{-1} = \frac{2a+1}{n-1} \xrightarrow{n \rightarrow \infty} 0, \quad (4.59)$$

dado que o valor máximo de a é W_{q^*} , ou seja, este é limitado. Agora, veja que, para dois conjuntos de números $a_i \geq 0$ e $b_i > 0 \forall i = 1, 2, \dots, n$, devemos ter:

$$\begin{aligned} \frac{a_i}{b_i} &\leq \max_{1 \leq i \leq n} \left(\frac{a_i}{b_i} \right) = \frac{a_*}{b_*} \Rightarrow a_i \leq \frac{a_*}{b_*} b_i \\ \therefore \sum_i a_i &\leq \sum_{i=1}^n \frac{a_*}{b_*} b_i = \frac{a_*}{b_*} \sum_{i=1}^n b_i \Rightarrow \frac{\sum_i a_i}{\sum_{i=1}^n b_i} \leq \frac{a_*}{b_*} = \max_{1 \leq i \leq n} \left(\frac{a_i}{b_i} \right). \end{aligned} \quad (4.60)$$

Note que a razão entre os termos provenientes de $(\mathbb{B}^{n-1})_{i,k}^2$ e de $\nu_{i,k}^2[n]$ sempre tende a zero. Portanto, como acabamos de demonstrar, o limite de $(\mathbb{B}^{n-1})_{i,k}^2 / \nu_{i,k}^2[n]$ também deve tender a zero para $n \rightarrow \infty$, e então a distribuição limite é gaussiana, como poderíamos esperar considerando as seções anteriores.

Portanto, concluímos que, sendo a matriz \mathbb{B} diagonalizável ou não, a gaussianidade de $S_{i,k}$ dependerá do maior módulo entre os autovalores que contribuem para a soma, dada a dependência sobre os termos de \mathbb{P} e \mathbb{P}^{-1} , que podem anular o efeito do autovalor associado. Considerando uma matriz com autovalores cujo autovalor de maior módulo que contribui para a soma é λ , e que o ruído não apresenta distribuição normal, temos:

- Se $0 < |\lambda| \leq 1$ a distribuição limite de $S_{i,k}$ não será gaussiana, e não apresentará difusão, tendo a variância saturada;
- Se $|\lambda| \geq 1$ a distribuição limite de $S_{i,k}$ não será gaussiana, e apresentará difusão geométrica;
- Se $|\lambda| = 1$ a distribuição limite de $S_{i,k}$ será gaussiana, e apresentará difusão normal, ao menos para $n \gg 1$.

A partir disto, pode-se avaliar a distribuição limite das componentes x_i , somando $S_{i,k} \forall k$. Sabendo que a soma de variáveis gaussianas também deve ter variáveis gaussianas, se as distribuições limites de $S_{i,k}$ forem gaussianas para todos os pares (i, k) – o que é mais provável se todos os autovalores de \mathbb{B} tiverem módulo unitário –, então todas as componentes de x_i terão distribuição limite gaussiana, ou seja, o sistema como um todo terá como distribuição limite uma gaussiana multidimensional.

5 Exemplos

5.1 A Equação de Langevin

A equação de Langevin é um modelo para o movimento de uma partícula imersa num fluido, considerando as colisões com as moléculas do fluido, assim como o atrito devido á viscosidade do fluido:

$$m \frac{d^2 x(t)}{dt^2} = -b \frac{dx(t)}{dt} + f(t), \quad (5.1)$$

sendo x a posição da nossa partícula, com b sendo o coeficiente de atrito e $f(t)$ a força aleatória causada pelas colisões com as moléculas do fluido, que consideramos que seja uma variável aleatória centralizada, estacionária e sem autocorrelação temporal. Primeiramente, transformamos o sistema em um de primeira ordem:

$$\frac{d\mathbf{x}(t)}{dt} = \mathbb{A}\mathbf{x}(t) + \mathbf{f}(t), \quad (5.2)$$

$$\mathbf{x}(t) = \begin{pmatrix} x_1(t) \\ x_2(t) \end{pmatrix} = \begin{pmatrix} x(t) \\ \frac{dx(t)}{dt} \end{pmatrix}; \quad \mathbf{f}(t) = \begin{pmatrix} 0 \\ f(t)/m \end{pmatrix}; \quad \mathbb{A} = \begin{pmatrix} 0 & 1 \\ 0 & -b/m \end{pmatrix}. \quad (5.3)$$

Note que \mathbb{A} não é inversível, e portanto não podemos utilizar a fórmula (12) para discretizar o sistema. Portanto, precisamos discretizar o sistema de outra maneira. Note que a evolução da velocidade x_2 não depende da posição x_1 , e podemos começar discretizando x_2 , tratando-o como um sistema unidimensional. Para cada intervalo no qual $f(t)$ é fixo em $f[n+1]$, temos:

$$x_2(t_n + t) = e^{-b(t-t_n)/m} x_2[n] + \frac{1 - e^{-b(t-t_n)/m}}{b} f[n+1], \quad t \leq t_{n+1} \quad (5.4)$$

e utilizamos isto para determinar a discretização de x_1 :

$$\int_{x_1[n]}^{x_1[n+1]} x_1 dx_1 = \int_{t_n}^{t_{n+1}} dt \left\{ e^{-b(t-t_n)/m} x_2[n] + \frac{1 - e^{-b(t-t_n)/m}}{b} f[n+1] \right\} \quad (5.5)$$

$$\begin{aligned}
x_1[n+1] - x_1[n] &= \frac{1 - e^{-b\Delta t/m}}{b/m} x_2[n] + \frac{1}{b} \left[\Delta t - \frac{1 - e^{-b\Delta t/m}}{b/m} \right] f[n+1] \\
&= \frac{m}{b} \left\{ x_2[n] - x_2[n+1] + \frac{\Delta t}{m} f[n+1] \right\}.
\end{aligned} \tag{5.6}$$

Isto nos permite escrever $x_1[n]$ muito facilmente em termos de $x_2[n]$:

$$x_1[n] = x_1[0] + \frac{m}{b} \left\{ x_2[0] - x_2[n] + \frac{\Delta t}{m} \sum_{j=1}^n f[j] \right\}. \tag{5.7}$$

Desta maneira, nos basta analisar a evolução de x_2 para conhecermos também a evolução de x_1 .

Vamos avaliar x_2 . Sua equação discretizada é dada por:

$$x_2[n+1] = Ax_2[n] + \eta[n+1], \tag{5.8}$$

$$A = e^{-b\Delta t/m}; \tag{5.9}$$

$$\eta[n] = \frac{1 - e^{-b\Delta t/m}}{b} f[n]. \tag{5.10}$$

Escrevendo o problema como uma soma de variáveis aleatórias:

$$x_2[n] = Ax_2[0] + S_2[n], \quad S_2[n] = \sum_{j=1}^n A^{n-j} \eta[j] = \sum_{j=0}^{n-1} A^j \eta[j]. \tag{5.11}$$

Sendo a função característica de $f[n]$ dada por

$$\psi_{\eta[n]}(z) = \psi_{\eta}(z) = \exp \left[-\frac{1}{2} M z^2 (1 + \omega_{\eta}(z)) \right], \tag{5.12}$$

$$M = \langle \eta[n]^2 \rangle = \left(\frac{1 - e^{-b\Delta t/m}}{b} \right)^2 \langle f[n]^2 \rangle, \tag{5.13}$$

escrevemos a função característica de $S_2[n]$:

$$\begin{aligned}
\psi_{S_2[n]}(z) &= \exp \left\{ -\frac{1}{2} M z^2 \sum_{j=0}^{n-1} A^{2j} [1 + \omega_{\eta}(A^j z)] \right\} \\
&= \exp \left\{ -\frac{1}{2} M^{(n)} z^2 [1 + \omega_{S_2[n]}(A^j z)] \right\},
\end{aligned} \tag{5.14}$$

$$M^{(n)} = M \sum_{j=0}^{n-1} A^{2j} = M \frac{A^{2n} - 1}{A^2 - 1}, \quad (5.15)$$

$$\omega_{S_2[n]} = \frac{1}{M^{(n)}} \sum_{j=0}^{n-1} A^{2j} \omega_\eta(A^j z). \quad (5.16)$$

Como $A < 1$, já sabemos que a distribuição limite não será gaussiana, mas deve se aproximar desta a medida que $b\Delta t/m$ tende a zero. Além de determinar a não gaussianidade da distribuição limite, podemos calcular explicitamente a evolução da variância, de modo a estimar a difusão do sistema. Já sabemos que esta será dada por $M^{(n)}$, e vamos escrevê-la por completo em termos dos parâmetros do sistema:

$$\langle S_2[n]^2 \rangle = M^{(n)} = \left(\frac{1 - e^{-b\Delta t/m}}{b} \right)^2 \frac{1 - e^{-2nb\Delta t/m}}{1 - e^{-2b\Delta t/m}}, \quad (5.17)$$

Lembrando que $n\Delta t = t$, isto é o tempo original, do sistema contínuo, temos:

$$\langle S_2(t)^2 \rangle = \left(\frac{1 - e^{-b\Delta t/m}}{b} \right)^2 \frac{1 - e^{-2bt/m}}{1 - e^{-2b\Delta t/m}} \xrightarrow{\Delta t \ll 1} \frac{1 - e^{-2bt/m}}{2bm} \Delta t \langle f^2 \rangle. \quad (5.18)$$

Para um tempo suficientemente longo, a variância de x_2 tende a $\Delta t \langle f^2 \rangle / (2bm)$, se $\Delta t \ll 1$, como é de se esperar, conhecendo a solução do movimento browniano sem discretização da equação (TOME; OLIVEIRA, 2015). Note que, se $\Delta t \rightarrow 0$, a variância se anula, e temos um sistema determinístico. Um artifício simples e comum para solucionar isso é embutir Δt em f , fazendo $f \rightarrow f/\sqrt{\Delta t}$, de modo a não se anular a contribuição do ruído quando se toma o limite $\Delta t \rightarrow 0$. Fazendo isso, é importante notar que a difusão da velocidade, que apresenta saturação na variância, acontece independente do intervalo Δt escolhido, mesmo que ele seja infinitesimal, ou seja, que consideremos a equação diferencial estocástica.

Agora, podemos partir para a análise de x_1 . Conhecendo $x_2[n]$, podemos reescrever e Equação (5.7):

$$x_1[n] = x_1[0] + \frac{m}{b} \left\{ \left(1 - e^{-nb\Delta t/m} \right) x_2[0] - S_1[n] \right\}, \quad (5.19)$$

$$S_1[n] = \sum_{j=0}^{n-1} \left[\frac{1 - e^{-b\Delta t/m}}{b} e^{-jb\Delta t/m} - \frac{\Delta t}{m} \right] f[j]. \quad (5.20)$$

A função característica de $S_1[n]$ será dada por

$$\psi_{S_1[n]}(z) = \exp \left\{ -\frac{1}{2} m^{(n)} z^2 \left[1 + \omega_{S_1[n]}(z) \right] \right\}, \quad (5.21)$$

$$m^{(n)} = m \sum_{j=0}^{n-1} A_1^2; \quad (5.22)$$

$$\omega_{S_1[n]} = \frac{1}{m^{(n)}} \sum_{j=0}^{n-1} A_1^2 \omega_f(A_1 z); \quad (5.23)$$

$$m = \langle f[n]^2 \rangle; \quad (5.24)$$

$$A_1 = \frac{1 - e^{-b\Delta t/m}}{b} e^{-jb\Delta t/m} - \frac{\Delta t}{m}. \quad (5.25)$$

Veja:

$$A_1^2 = \left(\frac{1 - e^{-b\Delta t/m}}{b} \right)^2 e^{-2jb\Delta t/m} - 2 \frac{\Delta t}{m} \frac{1 - e^{-b\Delta t/m}}{b} e^{-jb\Delta t/m} + \left(\frac{\Delta t}{m} \right)^2. \quad (5.26)$$

Note que o último termo não depende de j , e será somado linearmente, enquanto os outros se somarão em progressões geométricas de razões menores que 1 e, portanto, serão limitados. Portanto, a distribuição limite será gaussiana. Agora, vamos avaliar a evolução da variância que, conhecendo o movimento browniano, esperamos que, ao menos no limite $b\Delta t/m \rightarrow 0$, convirja para um comportamento linear.

$$\sum_{j=0}^{n-1} A_1^2 = \left(\frac{1 - e^{-b\Delta t/m}}{b} \right)^2 \frac{1 - e^{-2nb\Delta t/m}}{1 - e^{-2b\Delta t/m}} \quad (5.27)$$

$$- 2 \frac{\Delta t}{m} \frac{1 - e^{-b\Delta t/m}}{b} \frac{1 - e^{-nb\Delta t/m}}{1 - e^{-b\Delta t/m}} + \left(\frac{\Delta t}{m} \right)^2 n. \quad (5.28)$$

Com isto, agora conhecemos a variância de $S_1[n]$, que podemos escrever em termos do tempo t :

$$\langle S_1(t)^2 \rangle = \left\{ \left(\frac{1 - e^{-b\Delta t/m}}{b} \right)^2 \frac{1 - e^{-2bt/m}}{1 - e^{-2b\Delta t/m}} \right. \quad (5.29)$$

$$\left. - 2 \frac{\Delta t}{m} \frac{1 - e^{-b\Delta t/m}}{b} \frac{1 - e^{-bt/m}}{1 - e^{-b\Delta t/m}} + \frac{\Delta t}{m^2} t \right\} \langle f^2 \rangle \quad (5.30)$$

$$\xrightarrow{\Delta t \ll 1} \left\{ \frac{1 - e^{-2bt/m}}{2bm} - 2 \frac{1 - e^{-bt/m}}{bm} + \frac{t}{m^2} \right\} \Delta t \langle f^2 \rangle \quad (5.31)$$

Vemos que, de fato, o comportamento da variância tende a ser um comportamento linear a longo prazo, para $\Delta t \ll 1$. Lembrando, novamente, que poderíamos ter embutido Δt em f , de modo a manter a contribuição do ruído mesmo para $\Delta t \rightarrow 0$. Note que, na variância

de $x_1[n]$ ainda deve haver o fator $(m/b)^2$, enquanto $x_1[0]$ e $x_2[0]$ definem apenas o valor central da distribuição, caso estes sejam fixos, e não contribuem para a variância.

Podemos simular o sistema para confirmar as previsões aqui apresentadas. Escolhendo f como uma distribuição uniforme entre -1 e 1 e calculando 10000 realizações do sistema, cada uma com 1000 passos, obtemos a Figura 2, se escolhermos $b = 0.005$, $m = 1$, $\Delta t = 1$. Vemos que as distribuições de x_1 e x_2 são gaussianas neste caso. Isto é esperado para x_1 , e acontece também para x_2 devido à escala de $b\Delta t/m$, como já podíamos esperar.

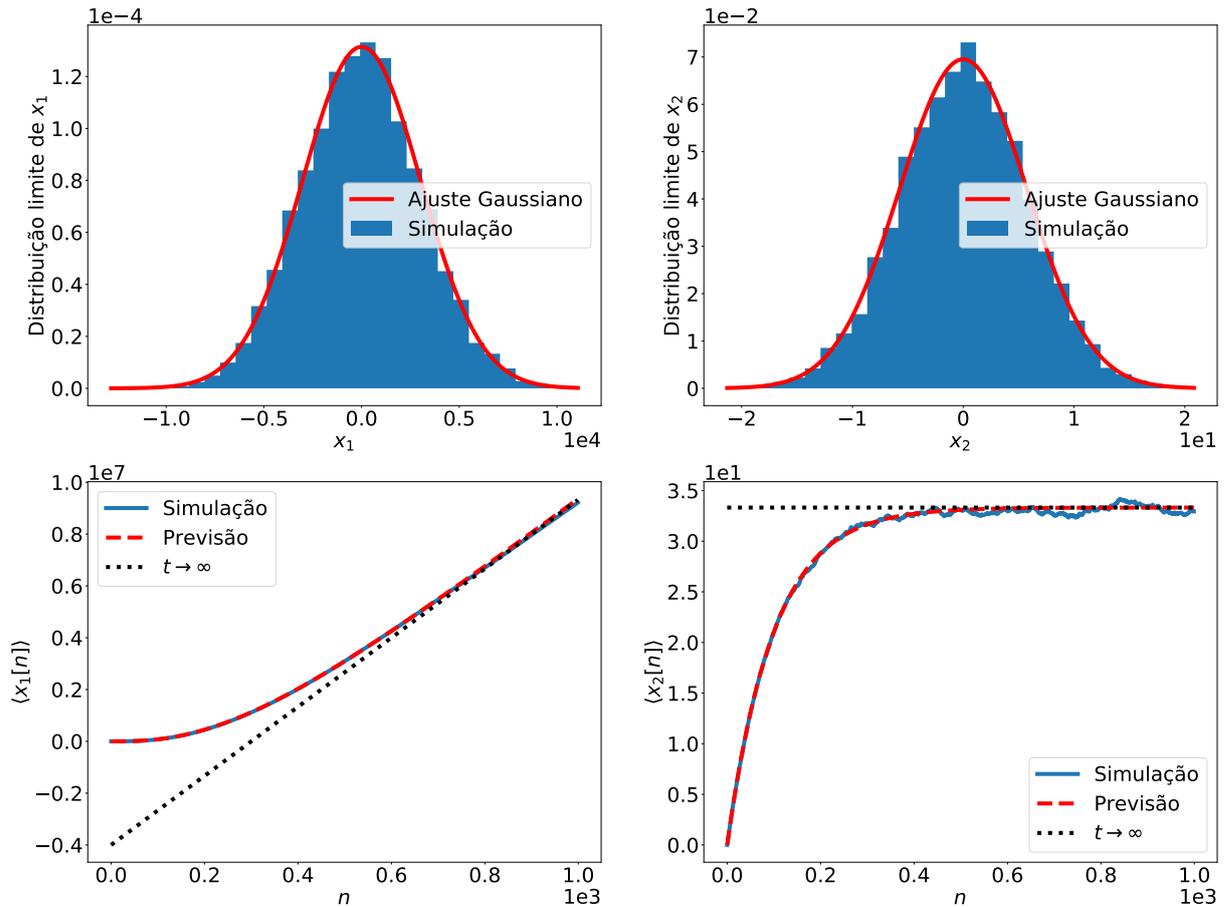


Figura 2 – Simulação da equação de Langevin para $b = 0.005$, $m = \Delta t = 1$, considerando 10000 trajetórias de 1000 passos.

Nos gráficos de cima, apresentamos em azul a distribuição da posição e da velocidade no último passo do sistema, considerando todas as 10000 realizações. Em vermelho, apresentamos a curva gaussiana que seria esperada para a distribuição, caso ela fosse gaussiana de fato, considerando a variância da distribuição obtida. Nos gráficos de baixo, temos a evolução das variâncias com o passar do tempo. Em azul, temos os resultados obtidos das simulações, enquanto em vermelho apresentamos os resultados esperados de acordo com os cálculos feitos acima, além da curva em preto, que retrata o comportamento esperado a longo prazo para a variância.

Na Figura 3, apresentamos os resultados para $b = 2$, mantendo m , Δt e a distribuição da força f . Note que, neste caso, as variâncias convergem muito mais rápido para seus valores esperados a longo prazo, se comparado à Figura 2.

O resultado mais importante na Figura 3 é que a distribuição limite da velocidade não é gaussiana se $b\Delta t/m$ não tende a zero. Porém, calcular diretamente a expressão para essa distribuição não é uma tarefa simples, como discutimos na seção 1.4.

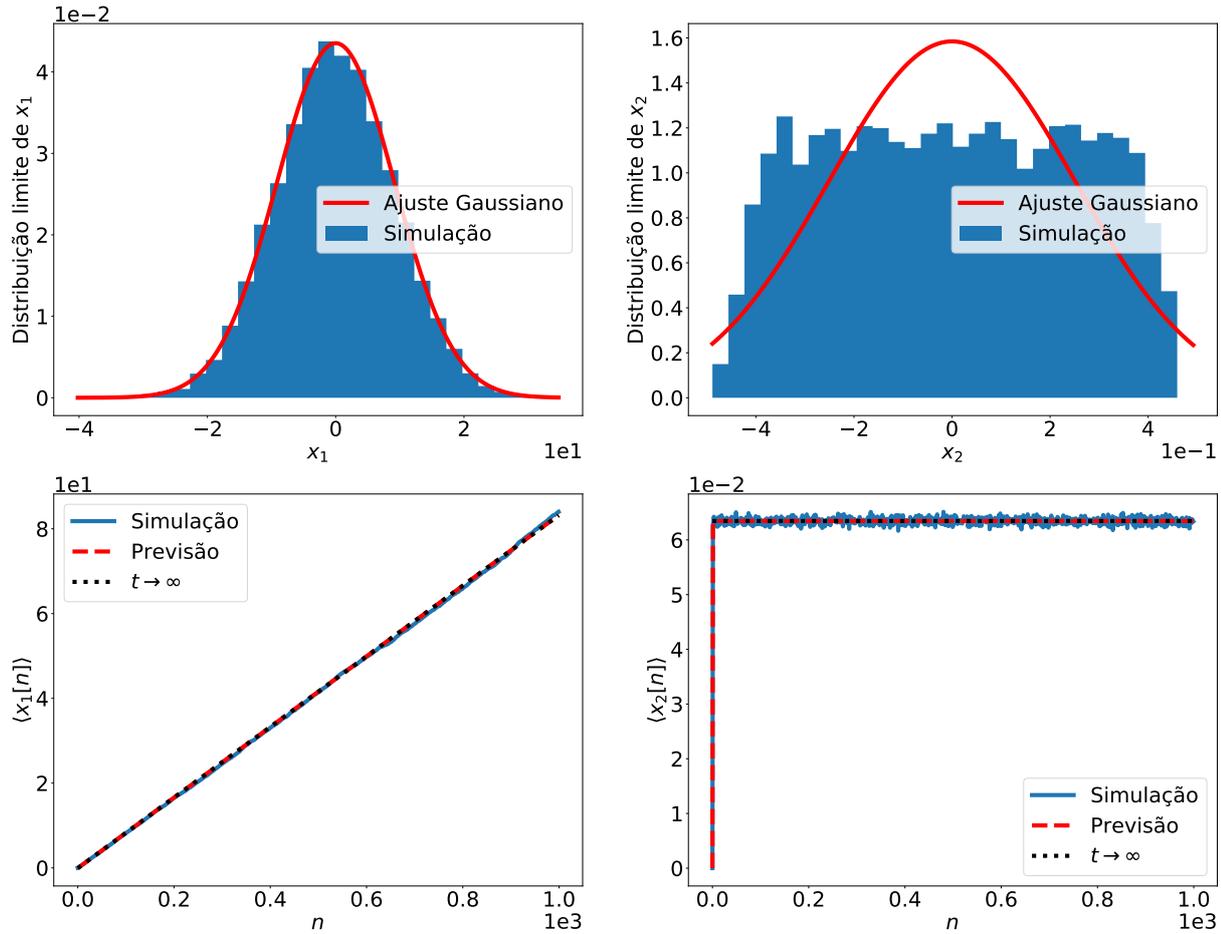


Figura 3 – Simulação da equação de Langevin para $b = 2$, $m = \Delta t = 1$, considerando 10000 trajetórias de 1000 passos.

5.2 O Oscilador Harmônico num Fluido

Agora, vamos avaliar um caso semelhante ao da seção anterior, de uma partícula num fluido, mas agora sob um potencial harmônico. Neste caso, a equação do sistema é dada por:

$$m \frac{d^2 x(t)}{dt^2} = -kx(t) - b \frac{dx(t)}{dt} + f(t), \quad (5.32)$$

sendo k o coeficiente elástico. Passando para um sistema de EDOs de primeira ordem, temos:

$$\frac{d\mathbf{x}(t)}{dt} = \mathbb{A}\mathbf{x}(t) + \mathbf{f}(t), \quad (5.33)$$

$$\mathbf{x}(t) = \begin{pmatrix} x_1(t) \\ x_2(t) \end{pmatrix} = \begin{pmatrix} x(t) \\ \frac{dx(t)}{dt} \end{pmatrix}; \quad \mathbf{f}(t) = \begin{pmatrix} 0 \\ f(t)/m \end{pmatrix}; \quad \mathbb{A} = \begin{pmatrix} 0 & 1 \\ -k/m & -b/m \end{pmatrix}. \quad (5.34)$$

Agora, precisamos discretizar o sistema:

$$\mathbf{x}[n+1] = \mathbb{B}\mathbf{x}[n] + \eta[n+1], \quad (5.35)$$

$$\mathbb{B} = e^{\mathbb{A}\Delta t}; \quad \eta = \mathbb{A}^{-1} (e^{\mathbb{A}\Delta t} - \mathbb{1}) \mathbf{f}. \quad (5.36)$$

Note que a inversa de \mathbb{A} é dada por:

$$\mathbb{A}^{-1} = \frac{m}{k} \begin{pmatrix} -b/m & -1 \\ k/m & 0 \end{pmatrix} = \begin{pmatrix} -b/k & -m/k \\ 1 & 0 \end{pmatrix} \quad (5.37)$$

Para calcular a exponencial de \mathbb{A} , precisamos antes diagonalizá-la. Seu polinômio característico é dado por:

$$\det(\mathbb{A} - \lambda\mathbb{1}) = \lambda \left(\lambda + \frac{b}{m} \right) + \frac{k}{m}. \quad (5.38)$$

Como estamos considerando um oscilador harmônico amortecido, temos três comportamentos possíveis, baseado nas relação entre b , k e m :

- Subcrítico ($b^2 < 4km$): O sistema apresenta oscilações com amplitude decaindo exponencialmente;
- Supercrítico ($b^2 > 4km$): O sistema decai exponencialmente, sem oscilar;
- Crítico ($b^2 = 4km$): O sistema decai, agora mais rapidamente do que no caso supercrítico.

Note que, dependendo do ruído η obtido, os ruídos nas duas componentes serão correlacionados entre si. Porém, isto não impede o cálculo de cada componente $S_{i,k}$. Vamos

considerar cada caso separadamente, começando pelo caso supercrítico. Neste caso, os autovalores de A serão dados por:

$$\lambda_{\pm} = -\frac{b}{2m} \pm \sqrt{\frac{b^2}{4m^2} - \frac{k}{m}}, \quad (5.39)$$

ambos negativos. Podemos escrever $\mathbb{B} = \mathbb{P}\mathbb{D}_{\mathbb{B}}\mathbb{P}^{-1}$, $\mathbb{D}_{\mathbb{B}} = \text{diag}(\lambda_-, \lambda_+)$. Calculando os autovetores de \mathbb{B} , determinamos a matriz \mathbb{P} :

$$\mathbb{P} = \begin{pmatrix} -\frac{b - \sqrt{b^2 - 4km}}{2k} & -\frac{b + \sqrt{b^2 - 4km}}{2k} \\ 1 & 1 \end{pmatrix}, \quad (5.40)$$

e, conseqüentemente, sua inversa:

$$\mathbb{P}^{-1} = \begin{pmatrix} \frac{k}{\sqrt{b^2 - 4km}} & \frac{1}{2} + \frac{b}{2\sqrt{b^2 - 4km}} \\ -\frac{k}{\sqrt{b^2 - 4km}} & \frac{1}{2} - \frac{b}{2\sqrt{b^2 - 4km}} \end{pmatrix}. \quad (5.41)$$

Calculando a exponencial de $A\Delta t$:

$$\begin{aligned} \mathbb{B} &= \frac{e^{-\frac{\Delta t(\sqrt{b^2 - 4km} + b)}{2m}}}{\sqrt{b^2 - 4km}} \times \\ &\times \begin{pmatrix} \frac{b \left(e^{\frac{\Delta t\sqrt{b^2 - 4km}}{m}} - 1 \right) + \sqrt{b^2 - 4km} \left(e^{\frac{\Delta t\sqrt{b^2 - 4km}}{m}} + 1 \right)}{2} & m \left(e^{\frac{\Delta t\sqrt{b^2 - 4km}}{m}} - 1 \right) \\ -k \left(e^{\frac{\Delta t\sqrt{b^2 - 4km}}{m}} - 1 \right) & b \left(1 - e^{\frac{\Delta t\sqrt{b^2 - 4km}}{m}} \right) + \sqrt{b^2 - 4km} \left(e^{\frac{\Delta t\sqrt{b^2 - 4km}}{m}} + 1 \right) \end{pmatrix}, \end{aligned} \quad (5.42)$$

podemos também calcular η :

$$\eta = \frac{f}{m} \begin{pmatrix} -\frac{me^{-\frac{\Delta t(\sqrt{b^2 - 4km} + b)}{2m}} \left(b \left(e^{\frac{\Delta t\sqrt{b^2 - 4km}}{m}} - 1 \right) + \sqrt{b^2 - 4km} \left(e^{\frac{\Delta t\sqrt{b^2 - 4km}}{m}} - 2e^{\frac{\Delta t(\sqrt{b^2 - 4km} + b)}{2m}} + 1 \right) \right)}{2k\sqrt{b^2 - 4km}} \\ \frac{me^{-\frac{\Delta t(\sqrt{b^2 - 4km} + b)}{2m}} \left(e^{\frac{\Delta t\sqrt{b^2 - 4km}}{m}} - 1 \right)}{\sqrt{b^2 - 4km}} \end{pmatrix}. \quad (5.43)$$

Percebe-se que as duas componentes estão diretamente correlacionadas, mas ainda podemos calcular as funções características das somas $S_{i,k}$. Porém, como as componentes

de η são correlacionadas, é difícil avaliar o comportamento de \mathbf{x} a partir das somas $S_{i,k}$. Uma possibilidade é passar para a base do vetor principal do ruído, que será nulo em uma das componentes. Porém, isso envolve multiplicar mais e mais matrizes tornando as expressões desnecessariamente grandes. Uma alternativa válida é avaliar diretamente a soma para cada componente do sistema, na base original.

Sabemos que podemos escrever:

$$\mathbf{x}[n] = \mathbb{B}^n \mathbf{x}[0] + \sum_{j=1}^n \mathbb{B}^{n-j} \eta[j] = \mathbb{B}^n \mathbf{x}[0] + \mathbf{S}[n], \quad \mathbf{S}[n] = \sum_{j=0}^{n-1} \mathbb{B}^j \eta[j], \quad (5.44)$$

de modo que, para cada componente de $\mathbf{S}[n]$, temos:

$$S_i[n] = \sum_{j=0}^{n-1} \left(\mathbb{B}^j \eta[j] \right)_i = \sum_{j=0}^{n-1} \sum_{k=1}^2 \left(\mathbb{B}^j \right)_{i,k} \eta_k[j]. \quad (5.45)$$

Se escrevemos a Equação (5.43):

$$\eta = \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} \zeta, \quad (5.46)$$

podemos reescrever:

$$S_i[n] = \sum_{j=0}^{n-1} \left[a_1 \left(\mathbb{B}^j \right)_{i,1} + a_2 \left(\mathbb{B}^j \right)_{i,2} \right] \zeta[j] = \sum_{j=0}^{n-1} c_{i,j} \zeta[j], \quad (5.47)$$

de modo que sua função característica é dada por:

$$\psi_{S_i[n]} = \exp \left\{ -\frac{z^2}{2} \sum_{j=0}^{n-1} \left[c_{i,j}^2 (1 + \omega_\zeta(c_{i,j}z)) \right] \right\}. \quad (5.48)$$

Note que podemos escrever:

$$\begin{aligned} c_{i,j}^2 &= \left(a_1 \left(\mathbb{B}^j \right)_{i,1} + a_2 \left(\mathbb{B}^j \right)_{i,2} \right)^2 = \left(a_1 \sum_{m=1}^2 \mathbb{P}_{i,m} \mathbb{P}_{m,1}^{-1} e^{j\lambda_m \Delta t} + a_2 \sum_{m=1}^2 \mathbb{P}_{i,m} \mathbb{P}_{m,2}^{-1} e^{j\lambda_m \Delta t} \right)^2 \\ &= \left(\sum_{m=1}^2 \mathbb{P}_{i,m} \left(a_1 \mathbb{P}_{m,1}^{-1} + a_2 \mathbb{P}_{m,2}^{-1} \right) e^{j\lambda_m \Delta t} \right)^2. \end{aligned} \quad (5.49)$$

$$\sum_{j=0}^{n-1} c_{i,j}^2 = \sum_{j=0}^{n-1} \left[\left(\mathbb{P}_{i,1} \left(a_1 \mathbb{P}_{1,1}^{-1} + a_2 \mathbb{P}_{1,2}^{-1} \right) \right)^2 e^{2j\lambda_1 \Delta t} + \left(\mathbb{P}_{i,2} \left(a_1 \mathbb{P}_{2,1}^{-1} + a_2 \mathbb{P}_{2,2}^{-1} \right) \right)^2 e^{2j\lambda_2 \Delta t} + \right. \\ \left. + 2 \mathbb{P}_{i,1} \mathbb{P}_{i,2} \left(a_1 \mathbb{P}_{1,1}^{-1} + a_2 \mathbb{P}_{1,2}^{-1} \right) \left(a_1 \mathbb{P}_{2,1}^{-1} + a_2 \mathbb{P}_{2,2}^{-1} \right) e^{j(\lambda_1 + \lambda_2) \Delta t} \right] \quad (5.50)$$

$$\sum_{j=0}^{n-1} c_{i,j}^2 = \left(\mathbb{P}_{i,1} \left(a_1 \mathbb{P}_{1,1}^{-1} + a_2 \mathbb{P}_{1,2}^{-1} \right) \right)^2 \frac{1 - e^{2n\lambda_1 \Delta t}}{1 - e^{2\lambda_1 \Delta t}} + \left(\mathbb{P}_{i,2} \left(a_1 \mathbb{P}_{2,1}^{-1} + a_2 \mathbb{P}_{2,2}^{-1} \right) \right)^2 \frac{1 - e^{2n\lambda_2 \Delta t}}{1 - e^{2\lambda_2 \Delta t}} + \\ + \mathbb{P}_{i,1} \mathbb{P}_{i,2} \left(a_1 \mathbb{P}_{1,1}^{-1} + a_2 \mathbb{P}_{1,2}^{-1} \right) \left(a_1 \mathbb{P}_{2,1}^{-1} + a_2 \mathbb{P}_{2,2}^{-1} \right) \frac{1 - e^{n(\lambda_1 + \lambda_2) \Delta t}}{1 - e^{(\lambda_1 + \lambda_2) \Delta t}} \quad (5.51)$$

$$\xrightarrow{n \rightarrow \infty} \frac{\left(\mathbb{P}_{i,1} \left(a_1 \mathbb{P}_{1,1}^{-1} + a_2 \mathbb{P}_{1,2}^{-1} \right) \right)^2}{1 - e^{2\lambda_1 \Delta t}} + \frac{\left(\mathbb{P}_{i,2} \left(a_1 \mathbb{P}_{2,1}^{-1} + a_2 \mathbb{P}_{2,2}^{-1} \right) \right)^2}{1 - e^{2\lambda_2 \Delta t}} + \\ + 2 \frac{\mathbb{P}_{i,1} \mathbb{P}_{i,2} \left(a_1 \mathbb{P}_{1,1}^{-1} + a_2 \mathbb{P}_{1,2}^{-1} \right) \left(a_1 \mathbb{P}_{2,1}^{-1} + a_2 \mathbb{P}_{2,2}^{-1} \right)}{1 - e^{(\lambda_1 + \lambda_2) \Delta t}}. \quad (5.52)$$

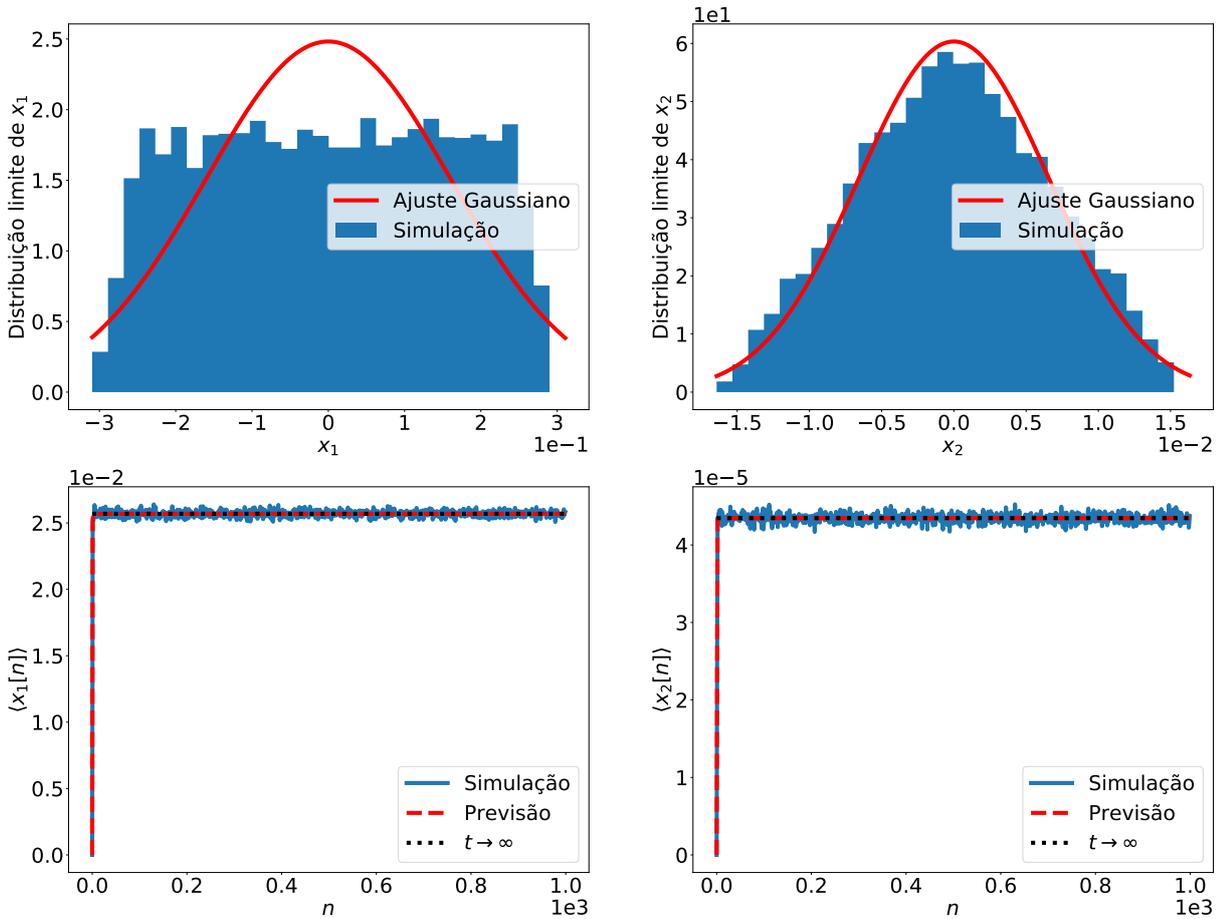


Figura 4 – Oscilador Harmônico num fluido. $b = 5$, $k = m = 1$: Regime supercrítico. $\Delta t = 10$.

Com isto, calculamos a evolução da variância para as duas componentes do sistema. Vale notar que esta expressão será a mesma também para os casos crítico e subcrítico, se $b > 0$, é claro, mantendo o amortecimento do sistema. Não há possibilidade de os autovalores se anularem para gerar um termo com crescimento linear no somatório. Portanto, como todos os termos apresentam crescimento limitado, as duas componentes devem apresentar variâncias limitadas, podendo também apresentar distribuições não gaussianas.

Vale notar que, se ao menos um dos expoentes envolvendo os autovalores se aproxima de zero, o termo envolvendo este expoente cresce rapidamente, tornando as contribuições das funções ω desprezíveis, e, conseqüentemente, tornando as distribuições gaussianas. Nas Figuras 4 e 5, temos alguns exemplos de possíveis comportamentos para o sistema, lembrando que x_1 e x_2 representam, respectivamente, a posição e a velocidade da partícula em questão. Novamente, f tem distribuição uniforme entre -1 e 1, e as simulações foram realizadas com 10000 trajetórias de 1000 passos.

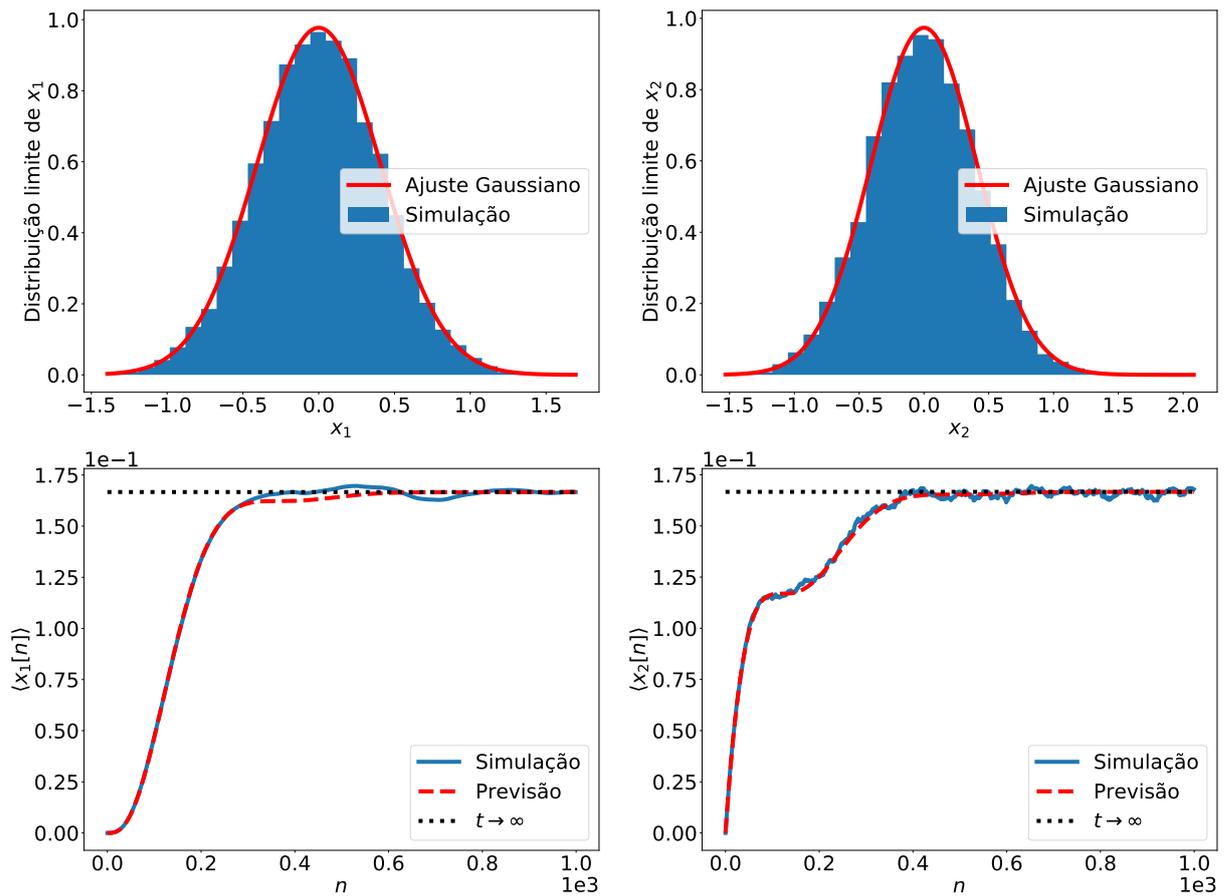


Figura 5 – Oscilador Harmônico num fluido. $b = k = m = 1$: Regime subcrítico. $\Delta t = 0.01$.

Note que, escolhendo um menor valor para Δt na Figura 5, conseguimos observar melhor o comportamento transiente na variância, que é especialmente interessante para a

velocidade, que apresenta um breve platô, antes de retomar a subida até seu valor limite. Vale ainda notar que, se tomamos o limite $b \rightarrow 0$, temos um oscilador sem amortecimento, para o qual temos apenas um comportamento possível, análogo ao subcrítico, agora com uma frequência de oscilação $\omega = \sqrt{k/m}$. Neste caso, os autovalores de \mathbb{A} serão $\lambda_{\pm} = \pm i\omega$, de modo que os autovalores de \mathbb{B} terão módulo unitário, e o terceiro termo da Equação (5.50) terá o produto dos autovalores anulando suas oscilações, e gerando um termo que terá crescimento linear com o somatório. Portanto as distribuições limites serão gaussianas, com variâncias crescentes com o passar do tempo, como vemos na Figura 6. Note o caráter oscilatório das variâncias.

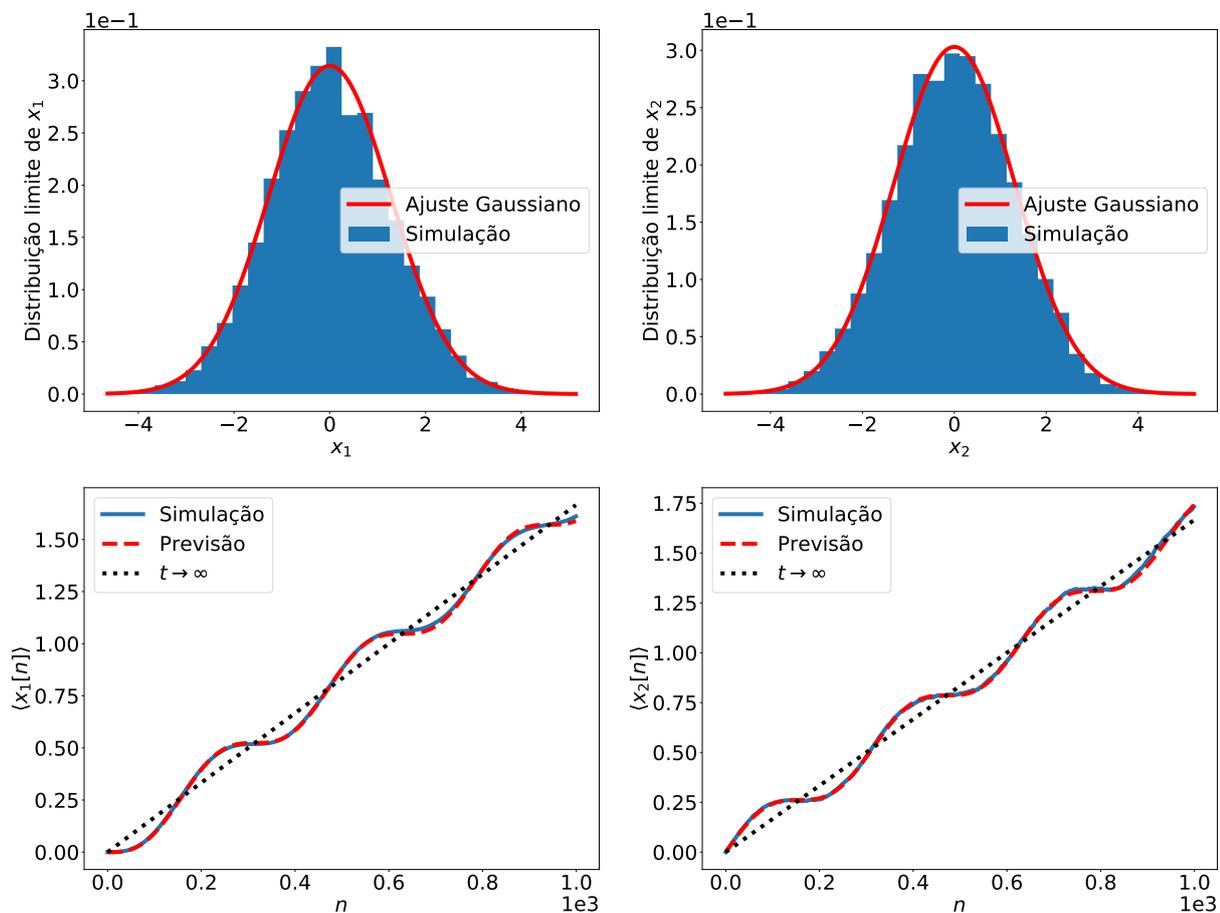


Figura 6 – Oscilador Harmônico com força aleatória, sem amortecimento. $k = m = 1$, $\Delta t = 0.01$

Conclusão

Este trabalho nos mostra que, ao considerarmos equações diferenciais lineares estocásticas com ruído constante por intervalos finitos, permitimos uma maior diversidade de comportamentos para o sistema, exemplificada pela não adequação ao teorema do limite central, apresentando distribuições limites não gaussianas.

Mostramos que é possível determinar, no caso em que as componentes do vetor aleatório principal são independentes entre si, as condições necessárias para a gaussianização da distribuição limite de cada componente do mapa vetorial e, a partir disto, sabemos quando o sistema como um todo será, no limite, determinado por uma distribuição normal multivariada. Utilizamos esta análise como base para estudar a equação de Langevin com força aleatória constante por intervalos finitos, e observamos alguns resultados bastante semelhantes àqueles da equação diferencial estocástica padrão. Para a posição, observamos distribuição limite gaussiana, e difusão normal; para a velocidade, obtemos saturação na variância. Porém, mostramos que a distribuição limite para a velocidade não é necessariamente gaussiana.

Este trabalho pode servir de motivação para, no caso de se utilizar um modelo linear com ruído aditivo, optar por um modelo de tempo discreto (caso isto faça sentido físico), no qual a distribuição do ruído original afeta a distribuição limite, caso não seja gaussiana.

Vale notar que, apesar de o desenvolvimento apresentado no Capítulo 4 se apresentar útil para a análise de sistemas de tempo discreto, pode não ser o melhor caminho para considerar sistemas de tempo contínuo através de sua discretização, uma vez que, como vimos no Capítulo 5, a mera discretização do sistema pode ser suficiente para tornar as componentes do ruído correlacionadas entre si, complicando a análise. Eventualmente, pode ser que as componentes do ruído se tornem independentes na base do vetor aleatório principal, mas isso deve envolver uma grande complicação dos cálculos, dadas as novas multiplicações de matrizes aumentando o tamanho das expressões. Tendo isso em mente, pode ser interessante o desenvolvimento de um pacote para sistemas de álgebra computacional que realize todos os cálculos a partir do sistema contínuo original.

É importante notar que, apesar de ser possível determinar se a distribuição limite será gaussiana ou não, em caso negativo, determinar qual será a distribuição limite de fato pode não ser uma tarefa simples. De fato, seria interessante continuar este estudo na direção de avaliar a precisão necessária na função ω do ruído original para determinação da função Ω da função característica limite com precisão suficiente para realização de previsões consistentes sobre a distribuição limite do sistema.

Ainda sobre possíveis estudos para dar continuidade a este trabalho, vale notar que a equação com que trabalhamos é um modelo de auto-regressão de ordem 1, ou seja, que estado em um instante depende apenas do instante anterior. Seria interessante considerar modelos auto-regressivos de maior ordem e determinar se, a medida que aumentamos a ordem do modelo auto-regressivo, as condições para que a distribuição limite seja gaussiana se tornam mais restritas ou mais abrangentes.

Referências

- BLACK, F.; SCHOLLES, M. The Pricing of Options and Corporate Liabilities. *Journal of Political Economy*, v. 81, n. 3, 5 1973. ISSN 0022-3808. Citado na página 17.
- BRONSON, R. *Matrix methods : an introduction*. [S.l.]: Academic Press, 1991. 503 p. ISBN 012135251X. Citado 2 vezes nas páginas 40 e 42.
- FIGUEIREDO, A. et al. Characteristic function approach to the sum of stochastic variables. *MPRA: Munich Personal RePEc Archive*, 2006. Citado na página 33.
- FRISTEDT, B.; GRAY, L. *A Modern Approach to Probability Theory*. Boston, MA: Birkhäuser Boston, 1997. ISBN 978-1-4899-2839-9. Citado na página 92.
- GNEDENKO, B. V.; KOLMOGOROV, A. N. *Limit Distributions for Sums of Independent Random Variables. Translated from the Russian by K. L. Chung. With an Appendix by J. L. Doob*. 2. ed. [S.l.]: Addison-Wesley Publishing Company, Cambridge, 1968. ISBN 13: 9780201024203. Citado na página 43.
- HARTMAN, P. On the local linearization of differential equations. *Proceedings of the American Mathematical Society*, American Mathematical Society (AMS), v. 14, n. 4, p. 568–568, 4 1963. ISSN 0002-9939. Citado na página 20.
- LÉVY, P. *Théorie des erreurs. La loi de Gauss et les lois exceptionnelles*. [S.l.: s.n.], 1924. Citado 2 vezes nas páginas 20 e 26.
- MARTINEZ-GARCIA, R. et al. How range residency and long-range perception change encounter rates. *Journal of Theoretical Biology*, Academic Press, v. 498, p. 110267, 8 2020. ISSN 10958541. Citado na página 17.
- ØKSENDAL, B. *Stochastic Differential Equations: An Introduction with Applications*. [S.l.]: Springer Science & Business Media, 2003. 360 p. ISBN 3540047581. Citado 2 vezes nas páginas 17 e 18.
- PALMER, T. N. Stochastic weather and climate models. *Nature Reviews Physics*, Nature Publishing Group, v. 1, n. 7, p. 463–471, 7 2019. ISSN 2522-5820. Citado na página 17.
- SHINOZUKA, M.; DEODATIS, G. Stochastic process models for earthquake ground motion. *Probabilistic Engineering Mechanics*, Elsevier, v. 3, n. 3, p. 114–123, 9 1988. ISSN 0266-8920. Citado na página 17.
- TOME, T.; OLIVEIRA, M. J. d. M. J. d. *Stochastic dynamics and irreversibility*. [S.l.]: Springer, 2015. ISBN 9783319117706. Citado 2 vezes nas páginas 23 e 75.

Apêndices

APÊNDICE A – Teoremas e Demonstrações

Teorema A.1 (1.1). *Considere um conjunto de variáveis aleatórias x_i , $i \in 1, 2, \dots, n$, sendo m_i o desvio padrão de x_i e M_n o desvio padrão da soma $X_n = x_1 + x_2 + \dots + x_n$. Se:*

1. $\exists \lim_{n \rightarrow \infty} (m_i/M_n) \forall i$;
2. $\exists L_z, |\omega_i(z')| \leq L_z, \forall i, \forall z', -z < z' < z$;
3. $\forall i, \omega_i(z)$ é contínua em $z = 0$;

então:

$$\exists \lim_{n \rightarrow \infty} \Psi_n^{(r)}(z) = \Psi^{(r)}(z) = \exp \left[-\frac{z^2}{2} (1 + \Omega(z)) \right],$$

onde $\Psi^{(r)}(z)$ é a função característica da variável reduzida da soma $X_n^{(r)}$.

Demonstração: Lembremos da Equações (1.65):

$$\Psi_n^{(r)}(z) = \prod_{i=1}^n \psi_i^{(r)} \left(\frac{m_i}{M_n} z \right),$$

e (1.18):

$$\psi_i^{(r)}(z) = \exp \left[-\frac{z^2}{2} (1 + \omega_i(z)) \right].$$

Com isto, podemos escrever:

$$\begin{aligned} \Psi^{(r)}(z) &= \prod_{i=1}^n \exp \left[-\frac{1}{2} \frac{m_i^2}{M_n^2} z^2 \left(1 + \omega_i \left(\frac{m_i}{M_n} z \right) \right) \right] = \exp \left\{ -\frac{z^2}{2} \sum_{i=1}^n \left[\frac{m_i^2}{M_n^2} \left(1 + \omega_i \left(\frac{m_i}{M_n} z \right) \right) \right] \right\} \\ \therefore \Psi^{(r)}(z) &= \exp \left[-\frac{z^2}{2} (1 + \Omega_n(z)) \right], \end{aligned} \tag{A.1}$$

onde $\Omega_n(z)$ é:

$$\Omega_n(z) = \sum_{i=1}^n \left[\frac{m_i^2}{M_n^2} \omega_i \left(\frac{m_i}{M_n} z \right) \right]. \tag{A.2}$$

Considerando a hipótese 1 do Teorema, podemos tomar o limite $n \rightarrow \infty$ de $\Omega_n(z)$:

$$\Omega(z) = \lim_{n \rightarrow \infty} \Omega_n(z) = \sum_{i=1}^{\infty} \lambda_i^2 \omega_i(\lambda_i z); \quad \lambda_i = \lim_{n \rightarrow \infty} \frac{m_i}{M_n}. \quad (\text{A.3})$$

A hipótese 2 do Teorema é responsável por garantir a convergência do limite de Ω_n . Note que:

$$\sum_{i=1}^{\infty} \lambda_i^2 |\omega_i(\lambda_i z)| \leq L_z \sum_{i=1}^{\infty} \lambda_i^2 = L_z \frac{1}{M_n^2} \sum_{i=1}^n m_i^2 = L_z. \quad (\text{A.4})$$

Portanto, pelo teste M de Weierstrass, podemos concluir que $\Omega(z)$ converge, dadas as hipóteses do Teorema.

A hipótese 3 do Teorema nos garante que $\Omega(z)$ é contínua em $z = 0$, dada sua definição (1.68). Então, utilizando o Teorema da Continuidade de Levi (apresentado e demonstrado na seção 14.7 de (FRISTEDT; GRAY, 1997)), concluímos que $\Psi^{(r)}(z)$ é uma função característica bem definida.

□