

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL
ESCOLA DE ENGENHARIA
INSTITUTO DE FÍSICA

DIONATAN CRISTIANO DA SILVA

**CRISES FINANCEIRAS DE CRIPTOATIVOS MODELADAS
ATRAVÉS DE UM MODELO BIESTÁVEL DA TEORIA DE
CATÁSTROFES**

TRABALHO DE DIPLOMAÇÃO EM ENGENHARIA FÍSICA II

PORTO ALEGRE
2019

DIONATAN CRISTIANO DA SILVA

**CRISES FINANCEIRAS DE CRIPTOATIVOS MODELADAS
ATRAVÉS DE UM MODELO BIESTÁVEL DA TEORIA DE
CATÁSTROFES**

Trabalho de Diplomação em Engenharia Física II apresentado à comissão de Graduação de Engenharia Física da Universidade Federal do Rio Grande do Sul, como requisito parcial para a obtenção do título de Bacharel em Engenharia Física.

Orientador: Prof. Dr. Carlo Requião da Cunha

PORTO ALEGRE
2019

AGRADECIMENTOS

Aos meus pais, Juvelino da Silva e Noeli da Silva, por terem me proporcionado, de diversas e indescritíveis formas, que eu chegasse até onde cheguei, permitindo-me oportunidades que a vida não lhes deu.

A minha namorada Cristiane da Rocha por inúmeras horas de apoio, incentivo e motivação que foram de extrema importância durante toda a minha caminhada.

Aos amigos que fiz durante meu período na universidade pelo companheirismo e amizade que se mostraram ser vitais. Em especial para Lucas Battu, Lucas Fialho, Franciele Mendes e Henrique Fonteles por todo apoio e parceria durante todos os anos de graduação.

Ao meu orientador Dr. Carlo Requião da Cunha, pela acolhida orientação não só nesse trabalho mas também no prosseguimento da vida acadêmica e profissional.

A minhas companheiras Thalia, Nove, Íris e Vênus que em muitos momentos me fizeram companhia dando suporte emocional durante o período de graduação.

À Universidade Federal do Rio Grande do Sul por me proporcionar oportunidades, aprendizados e experiências que levarei comigo eternamente.

O maior inimigo do conhecimento não é a ignorância, é a ilusão do conhecimento. (Stephen Hawking).

RESUMO

DA SILVA, Dionatan Cristiano. Crises financeiras de criptoativos modeladas através de um modelo biestável da teoria de catástrofes. 2019. 41 f. Trabalho de Diplomação em Engenharia Física II – Instituto de Física, Universidade Federal do Rio Grande do Sul. Porto Alegre, 2019.

No mercado econômico, os preços das criptomoedas costumam apresentar altas e baixas temporárias. Uma possível explicação para as mudanças abruptas observadas é oferecida pelo modelo de catástrofe.

A teoria de catástrofe começa na década de 70 com os trabalhos do francês René Thom e desde então tem sido utilizada em diversas áreas do conhecimento para descrever inúmeros eventos que acontecem de maneira repentina. São conhecidos alguns tipos de funções elementares de catástrofe e este trabalho traz um apanhado das funções de uma variável para a utilização no mercado financeiro.

Têm-se o objetivo analisar mercados de criptoativos para ajustes por um modelo biestável do tipo *cusp*. Foi visto que um ajuste a esse modelo pode trazer informações importantes a respeito do risco do mercado e pontos de instabilidade dos preços dos ativos.

Palavras-chave: Econofísica. Crises. Catástrofe. Cusp.

ABSTRACT

DA SILVA, Dionatan Cristiano. Cryptoactive financial crises modeled by a bistable model of catastrophe theory. 2019. 41 f. Trabalho de Diplomação em Engenharia Física II – Instituto de Física, Universidade Federal do Rio Grande do Sul. Porto Alegre, 2019.

In economic markets, cryptocurrencies prices often exhibit temporary booms and busts. A possible explanation for the observed abrupt changes is offered by the catastrophe models.

The theory of catastrophe begins in the 1970s with the works of the French Rene Thom and has since been used in several areas of knowledge to describe innumerable events that happen suddenly. Some types of elementary catastrophe functions are known, and this paper provides an overview of the functions of one variable for use in the financial market.

The objective is to analyze crypto markets for adjustments by a bistable cusp model. It has been seen that an adjustment to this model can provide important information regarding market risk and asset price instability points.

Keywords: Econophysics. Crises. Catastrophe. Cusp.

LISTA DE FIGURAS

Figura 1 – Superfície <i>cusp</i>	7
Figura 2 – Função potencial <i>cusp</i>	8
Figura 3 – Gráfico de pizza para parcela de mercado das principais criptomoedas	14
Figura 4 – Série histórica de preços de fechamento diários do Bitcoin	14
Figura 5 – Série histórica de retornos logarítmicos diários do Bitcoin	16
Figura 6 – Histograma dos retornos do Bitcoin	17
Figura 7 – Imagem representativa da distância interquartil <i>IQR</i>	17
Figura 8 – Série temporal dos preços normalizados das quatro criptomoedas analisadas, de forma que todas começassem o período custando U\$ 1,00.	18
Figura 9 – Histograma de retornos das quatro criptomoedas analisadas no ano de 2018.	18
Figura 10 – Comparação da distribuição cumulativa de probabilidade para os retornos das quatro moedas e comparação com uma distribuição normal.	19
Figura 11 – Gráfico do ajuste feito para a validação do algoritmo	25
Figura 12 – Série temporal de todos os dados trabalhados nas análises.	26
Figura 13 – <i>Cluster</i> realizado para agrupamento dos dados trabalhados.	29
Figura 14 – Ajuste da criptomoeda Bitcoin a um modelo biestável em dados de 2018	30
Figura 15 – Ajuste da criptomoeda Litecoin a um modelo biestável em dados de 2018	31
Figura 16 – Ajuste da criptomoeda Ethereum a um modelo biestável em dados de 2018	32
Figura 17 – Ajuste da criptomoeda Ripple a um modelo biestável em dados de 2018	33

LISTA DE ABREVIATURAS E SIGLAS

XRP	Criptomoeda Ripple
BTC	Criptomoeda Bitcoin
LTC	Criptomoeda Litecoin
ETH	Criptomoeda Ethereum
SOC	Criticalidade auto-organizada (Self Organized Criticality, do inglês)
FED	Série temporal para dados <i>overnight</i> da taxa de juros americana no ano de 2018
NVIDIA	Série temporal para os preços das ações da Nvidia durante o ano de 2018
AMD	Série temporal para os preços das ações da AMD durante o ano de 2018
S&P500	Série temporal para o valor do índice americano S&P500 durante o ano de 2018
M2	Série temporal de dados de quantidade de dólar circulante no mercado durante o ano de 2018
KS	Kolmogorov-Smirnov
CDF	Distribuição acumulada de probabilidades (Cumulative Distribution Function, do inglês)

SUMÁRIO

1 – INTRODUÇÃO	1
1.1 OBJETIVOS	2
1.2 ORGANIZAÇÃO DO TRABALHO	3
2 – REVISÃO BIBLIOGRÁFICA	4
2.1 INTRODUÇÃO À TEORIA DE CATÁSTROFES	4
2.1.1 MODELO DE BIESTABILIDADE	6
3 – O PROBLEMA	11
3.1 O QUE É UMA CRISE FINANCEIRA	11
3.2 TIPOS DE MOEDAS	11
3.3 CRIPTOMOEDAS	12
3.3.1 BITCOIN	12
3.3.2 BLOCKCHAIN	12
3.3.3 MINERAÇÃO DE CRIPTOMOEDAS	13
3.3.4 MERCADO DE CRIPTOMOEDAS	13
3.4 ABORDAGEM UTILIZADA PARA CRIPTOMOEDAS	16
3.4.1 APLICAÇÃO DO TESTE KS ÀS CRIPTOMOEDAS	20
4 – METODOLOGIA	21
4.1 PROJETO	21
4.1.1 IMPLEMENTAÇÃO DO CÓDIGO	22
4.1.1.1 AJUSTE	22
4.2 SELEÇÃO DE VARIÁVEIS	23
4.2.1 DIVISÃO EM <i>CLUSTERS</i>	23
4.3 ALGORITMO UTILIZADO	24
4.3.1 VALIDAÇÃO DO ALGORITMO	24
5 – ANÁLISE E DISCUSSÃO DOS RESULTADOS	26
5.1 DADOS	27
5.2 RESULTADOS OBTIDOS	27
5.2.1 CLUSTERS	28
5.2.2 RESULTADOS DOS AJUSTES	29
6 – CONCLUSÃO	34
6.1 TRABALHOS FUTUROS	34

Referências	35
Apêndices	37
APÊNDICE A—MÉTODO DE LEVENBERG–MARQUARDT	38
APÊNDICE B—TESTE DE KOLMOGOROV-SMIRNOV	40

1 INTRODUÇÃO

O mercado financeiro pode ser interpretado como o ambiente da economia onde se realizam todas as transações com moedas e títulos,¹ e as participações de capital². O comportamento imprevisível dos mercados de ações, especialmente crises inesperadas, tem sido um pesadelo para o mundo financeiro desde que os mercados de capital passaram a existir.

O matemático francês René Thom desenvolveu, no século passado, a teoria de catástrofes. Esta é uma teoria geométrica que tenta explicar porque algumas variações bruscas em sistemas aparecem na natureza. Este trabalho irá apresentar algumas ideias a respeito desta teoria.

Modelos matemáticos são, em sua maioria, construídos para descrever fenômenos que encontramos na natureza. Em alguns casos, quando se pode reproduzir um fenômeno por uma única equação, esse modelo pode até ser considerado uma lei da própria natureza. A física, como ciência, é repleta de modelos criados para entendermos diversos dos fenômenos. Estes podem ser dos mais diversos possíveis, como conservação de energia, troca de estado, dinâmica de corpos, entre outros tantos.

Um sistema (físico, mecânico, um modelo econômico ou social, etc) é, muitas vezes, modelado por equações diferenciais que representam, ou tentam representar, a evolução do sistema ao longo do tempo.

A ideia de sociofísica é antiga e data do século 19 ([SAINT-SIMON, 1803](#)), já o termo econofísica é mais recente aparecendo pela primeira vez em meados dos anos 1990. A última surge como uma resposta a aplicação prática da economia, muito mais ligada ao campo da macroeconomia, que é um ramo que estuda os fenômenos econômicos em escala global por meios estatísticos e matemáticos. A principal diferença é que ela não parte de pressupostos, de paradigmas, ou de teorias; propondo-se a utilizar apenas os dados reais da economia e com eles fazer diversos cálculos e métricas ou, em poucos casos, possíveis previsões utilizando os métodos já conhecidos da física.

A econofísica e a sociofísica, como áreas da física, podem ser definidas, de certa maneira, como abordagens quantitativas de fenômenos sócio-econômicos, aplicando ferramentas e conhecimentos de probabilidade e métodos estatísticos que geralmente são utilizados da física estatística.

Desde a década de 80 tem sido observado um crescente número de publicações utilizando física estatística para modelar fenômenos sociais dentre os quais alguns que se

¹Títulos financeiros são papéis que governos ou empresas vendem para conseguir os recursos financeiros que necessitam. É, basicamente, como se fossem empréstimos, onde quem recebe o dinheiro faz uma promessa de pagamento (com uma taxa acordada no título) a quem comprou o título.

²Podem ser entendidas como as ações de uma empresa. Ou seja, as parcelas mínimas que a companhia emite para venda em um mercado a fim de arrecadar recursos em troca dessa parcela da própria empresa. Uma vez essas parcelas sendo vendidas no primeiro momento pela empresa a um comprador, elas são vendidas e compradas entre outros compradores e vendedores no que é chamado mercado secundário.

destacam são formação de opinião, comportamento de grupo, evolução de linguagem e passagem de informação. Já a econofísica vem com o desafio de tentar entender a dinâmica de mercado com manifestações de correlação de longo alcance não lineares (KUTNER et al., 2019).

Resumindo, a física pode contribuir para a economia emprestando as ferramentas matemáticas que foram desenvolvidas ao longo dos anos para o estudo da matéria e da natureza como um todo (ARGYRAKIS, 2011). No desenvolvimento científico, a experiência mostra que as abordagens interdisciplinares são sempre úteis para o avanço do conhecimento.

Por outro lado, modelos matemáticos trazidos pela teoria da catástrofe podem contribuir para uma boa descrição dos dados empíricos de mercados sob mudanças bruscas nos preços, o que é uma característica do mercado de ações, por exemplo. Essas mudanças repentinas que ocorrem de várias formas são características internas e cruciais do sistema. Uma ideia-chave na teoria é que o sistema em estudo é direcionado para um estado de equilíbrio. De forma mais resumida, pode-se dizer que a evolução dos preços no mercado, sob a influência dos parâmetros de controle, geralmente ocorre de forma estável a longo prazo. Contudo, ocasionalmente esta evolução entra em áreas de instabilidade.

1.1 OBJETIVOS

Este trabalho tem como objetivo principal introduzir o estudo de teoria de catástrofe e mostrar a sua relevância no ramo de mercados financeiros. O trabalho terá como foco principal a abordagem em criptoativos, que são moedas virtuais ou também chamadas de criptomoedas, devido a sua maior volatilidade³ comparado ao mercado convencional e pela facilidade encontrada em obter dados. Normalmente a volatilidade é utilizada como uma medida de risco atrelada ao ativo econômico.

Visto isso, pode-se dizer que os três objetivos desse projeto são os seguintes:

- Aquisição de dados de vários ativos econômicos;
- Elaboração de algoritmo e código em Python para analisar os dados, verificar crises e tentar ajustá-las;
- Analisar os ajustes dos dados contendo crises através da teoria de catástrofe e tentar agrupá-los.

Cada um desses objetivos pode ser visto também como uma etapa do trabalho a ser seguida e ficará tudo mais claro conforme o progresso desta leitura.

³Volatilidade de um ativo econômico é uma medida de dispersão dos retornos desse ativo. Na literatura ela é geralmente introduzida como o desvio-padrão dos retornos desse ativo.

1.2 ORGANIZAÇÃO DO TRABALHO

Este trabalho de conclusão de curso está dividido em seis capítulos. No segundo capítulo, apresentamos uma revisão dos conceitos por trás da teoria, mostrando um pouco do contexto histórico dela, onde destacaremos os nomes dos grandes estudiosos e como iniciaram as pesquisas sobre este assunto.

Os capítulos seguintes são destinados a descrição do problema a ser abordado e a metodologia de estudo utilizada.

A análise dos resultados obtidos a partir dos modelos e hipóteses apresentadas durante a metodologia e as conclusões produzidas por esses resultados são exibidos na parte final do trabalho.

2 REVISÃO BIBLIOGRÁFICA

Esta seção tem como objetivo discutir brevemente as ideias básicas da teoria de catástrofes desde o seu surgimento no século passado até algumas aplicações mais recentes. Ainda faremos uma descrição um pouco mais detalhada do modelo que será utilizado para ajustar os dados de criptoativos.

A teoria, criada por René Thom, sempre atraiu um grande interesse por ser uma das primeiras teorias a explicar, ou pelo menos tentar explicar de forma mais incisiva, como uma mudança contínua de parâmetros do sistema pode causar uma variedade de fenômenos descontínuos.

2.1 INTRODUÇÃO À TEORIA DE CATÁSTROFES

Em geral, considera-se o nascimento da teoria de catástrofes como teoria matemática no ano de 1972, com a publicação de *Stabilité e Structurelleet Morphogénèse* (THOM, 1972), que foi o teste de uma teoria geral de modelos de R. Thom. A teoria de catástrofe é um ramo da teoria de sistemas dinâmicos que estuda e tenta classificar fenômenos caracterizados por mudanças repentinas no comportamento surgindo a partir de mudanças pequenas em alguns parâmetros.

A popularidade do tema aumentou muito com os trabalhos de Christopher Zeeman, no qual ele de forma brilhante exemplifica catástrofes através de um modelo de agressão¹ para raiva e medo em cachorros (ZEEMAN, 1976).

No modelo, o autor mostrou que os fatores que influenciam o comportamento de um cão com medo ou raiva são perfeitamente inferidos, respectivamente, pela inclinação das orelhas e da abertura da boca.

De acordo com o próprio Thom, o termo “Teoria de Catástrofes” foi criado por Zeeman. Zeeman também foi o primeiro a propor o uso de teoria de catástrofe para o comportamento do mercado de ações. Contudo, o número de artigos usando a teoria na economia ainda é razoavelmente pequeno.

Além do mercado de ações, vale a pena mencionar o mercado imobiliário, que é um outro exemplo de mercado que é um sistema dinâmico em que períodos de flutuações bruscas de preços podem ser definidos como momentos críticos em sua evolução temporal. Além disso, C. Zeeman, em seus trabalhos, considerou algumas aplicações populares do modelo de catástrofe em várias disciplinas diferentes, incluindo ecologia, física e psicologia.

Em sistemas dinâmicos, segundo M. Belej e S. Kulesza (BELEJ; KULESZA, 2012), pequenas perturbações podem deixar o sistema estruturalmente instável, pois estas variações nos parâmetros podem resultar em grandes mudanças estruturais do sistema como um todo.

¹Do inglês *Agression Model*.

Os autores acreditam que os métodos convencionais para análise de preços no mercado imobiliário, por exemplo, só são válidos em momentos de estabilidade enquanto que em períodos de instabilidade estes métodos se tornam insuficientes, dando lugar para uma possível nova linguagem para descrever tais períodos. É muito importante observar os períodos de instabilidade pois as flutuações não são apenas rupturas de tendências ou ciclos, mas sim partes essenciais do próprio mercado em questão.

Uma primeira premissa a ser assumida para tratar os eventos com o que a teoria de catástrofes diz é que o sistema seja do tipo gradiente, ou seja, para cada ponto tem-se um vetor que indica a direção na qual se tem o maior deslocamento do sistema. Isso quer dizer que existe uma função potencial a partir da qual podemos determinar os estados do sistema. Ou seja, uma função potencial é uma função de um certo número de variáveis, com essa função tendo valores reais.

Uma segunda premissa é a de negligenciar a parte transitória (escala de tempo muito curta) para que o foco seja nas posições de equilíbrio, instáveis ou estáveis. Ou seja, nos casos de potenciais de uma variável que serão avaliados nos próximos parágrafos é suficiente estudar a derivada da função potencial e igualá-la a zero para determinarmos as posições de equilíbrio.

Uma questão chave é verificar se o caminho para uma crise repentina é ou não descontínuo, e é por isso que teoria de catástrofes é comumente referida como teoria de mudanças descontínuas (STEWART, 1977). Uma forma diferente de colocarmos essa questão, é se saltos ou catástrofes de fato ocorrem e se eles podem ser de alguma forma evitados. É este tema e objetivo que muitos artigos têm em comum, de forma a analisar uma crise financeira do ponto de vista da teoria de catástrofe e muitos desses trabalhos conseguem resultados interessantes (WESSELBAUM, 2017).

René Thom desenvolveu uma maneira de observar a variação dos parâmetros em relação às variáveis dos modelos. Criou então o que ele denominou de famílias de catástrofes básicas chamadas de formas “canônicas” de catástrofes.

Thom também cita funções de mais de uma variável, o que totaliza sete diferentes famílias de modelos de catástrofe baseadas no número de variáveis de controle e de variáveis dependentes. Estas funções de mais de uma variável não cabem no escopo deste trabalho e por isso serão omitidas. Dentre os tipos de catástrofes de uma variável, existem quatro tipos de catástrofes conhecidos como catástrofes elementares de René Thom: *Fold*, *Cusp*, *Swallowtail* e *Butterfly*, sendo *Cusp* a mais comum delas.

Assim, como foi dito anteriormente e é observado em diversos tipos de mercados financeiros, pequenas mudanças dentro dos parâmetros de controle podem levar a mudanças significativas (estruturais) no sistema global, que é em grande parte o que a teoria de catástrofe quer englobar.

Outros artigos, como (BARUNIK; VOSVRDA, 2009), têm como principal objetivo responder se catástrofes são capazes de indicar *crashes* no mercado de ações. Neste artigo, é mostrado que o modelo de catástrofe *cusp* explica crises na bolsa de valores melhor do que

outros modelos. Os autores demonstram que a crise de 19/10/1987 nos Estados Unidos é melhor explicada pela catástrofe *cusp* e que para a crise de 11/09/2001 isso não é verdade. É interessante notar que a primeira crise é dada por fatores internos e a segunda por fatores externos.

Através da biestabilidade, modelar crises econômicas parece desafiador, mas relevante caso se consiga chegar a uma boa solução para este problema. Um bom exemplo é descrito por M. Belej e S. Kulesza (BELEJ; KULESZA, 2012). Nesse trabalho, os autores fizeram uma análise aprofundada de preços do mercado imobiliário da Polônia em termos de uma catástrofe do tipo *cusp*, que é o tipo elementar de catástrofe mais comumente utilizado na literatura e que também será utilizado nesse trabalho. Os resultados dos autores foram muito significativos, mostrando um melhor ajuste aos dados pelo modelo *cusp* frente a um modelo logístico.

Um outro exemplo muito interessante da literatura a respeito do tema é abordado em (DIKS; WANG, 2016), onde os autores observaram dados de seis países diferentes e descobriram evidências de que a dinâmica do mercado imobiliário desses países também pode ser explicada pelo comportamento biestável do tipo *cusp*.

Em métodos estatísticos clássicos, como a análise de regressão, as relações entre as variáveis são geralmente tratadas como perturbações de longo prazo e tendências estáveis. Também podemos dizer que o mercado evolui com o tempo sendo influenciado por muitos fatores. Alguns fatores são conhecidos, mas são tantas variáveis que fica impossível mapeá-las com exatidão. A teoria da catástrofe pode ser uma ferramenta útil para estimar o efeito qualitativo e até mesmo quantitativo de determinadas variáveis e suas combinações no comportamento de um sistema.

2.1.1 MODELO DE BIESTABILIDADE

Como dito anteriormente, o tipo *Cusp* é o mais comum dentre os modelos apresentados por R. Thom e dele serão mostradas algumas equações que, de maneira análoga, podem ser deduzidas para os outros tipos de catástrofes.

Considerando o potencial *Cusp* $V(x)$ do tipo:

$$V(x) = x^4 + ax^2 + bx . \quad (1)$$

A superfície de equilíbrio \mathcal{E} é dada por:

$$\mathcal{E} = \frac{dV}{dx} = 4x^3 + 2ax + b = 0 , \quad (2)$$

e o conjunto de singularidades \mathcal{S} (que ocorre quando $V_{xx} = 0$):

$$\frac{d^2V}{dx^2} = 12x^2 + 2a = 0 . \quad (3)$$

Dessa forma, apenas substituindo as equações 2 e 3 na equação do potencial para $V(x) = 0$, podemos concluir que o conjunto bifurcação, \mathcal{B} , da catástrofe *cusp* é da forma:

$$\mathcal{B} = 8a^3 + 27b^2 = 0 . \quad (4)$$

É graças a essa bifurcação que pode ser observado (para diferentes parâmetros a e b) se estamos tratando de regiões instáveis ou estáveis (Figura 1). Regiões instáveis tendem a estar dentro da bifurcação.

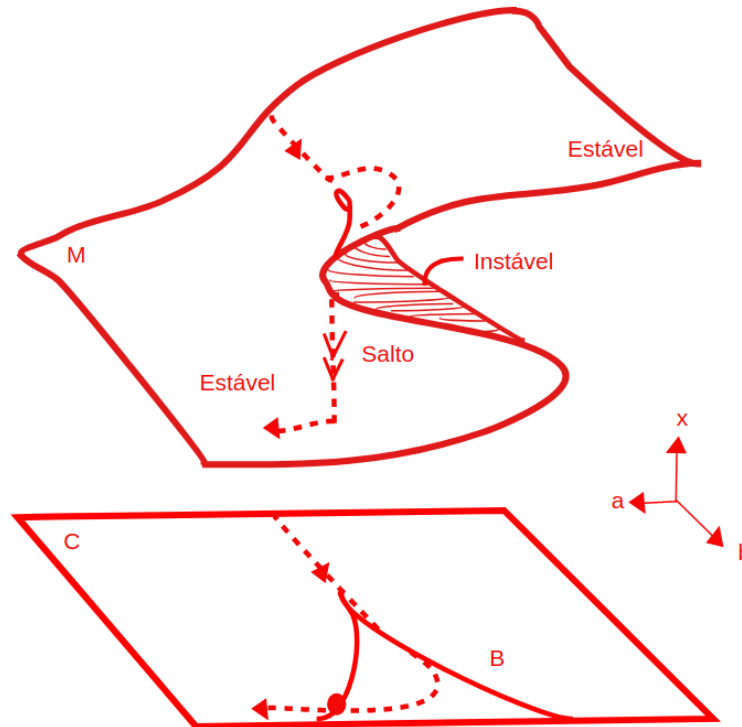


Figura 1 – Superfície do modelo de catástrofe *cusp*. Os parâmetros de controle são a e b . Adaptado de: (STEWART, 1977).

Os pontos estáveis e o ponto instável desse tipo de potencial para o tipo *cusp* são bem exemplificados na Figura 2 na qual está ilustrado o movimento de uma bola em uma superfície curva. Este sistema exibe três possíveis equilíbrios. Dois dos estados são estáveis, o que significa que o comportamento do sistema permanecerá relativamente inalterado quando houver uma perturbação. Um outro estado é instável, o que significa que uma pequena perturbação levará o sistema a um estado diferente.

A bimodalidade é outro aspecto importante do modelo *cusp*. Isso significa que, para um valor das variáveis de controle, existem dois possíveis comportamentos, da mesma forma que ocorre no modelo de Landau usado na mecânica estatística.

A catástrofe *fold* corresponde à perturbação de $y = x^3$ pela adição de uma função linear $y = x^3 + ax$ de modo a criar ($a < 0$) ou destruir ($a > 0$) um mínimo local. Da mesma forma, a catástrofe *cusp* corresponde à perturbação de $y = x^4$ pela adição de uma função quadrática $y = x^4 + ax^2 + bx$. Neste caso a e b são variáveis de controle que determinam o comportamento do sistema.

A Tabela 2.1.1 mostra alguns casos para diferentes valores de variáveis de controle. Pode-se perceber que se b é mantido fixo em 0 e a variar de -1 a $+1$, o mínimo local se

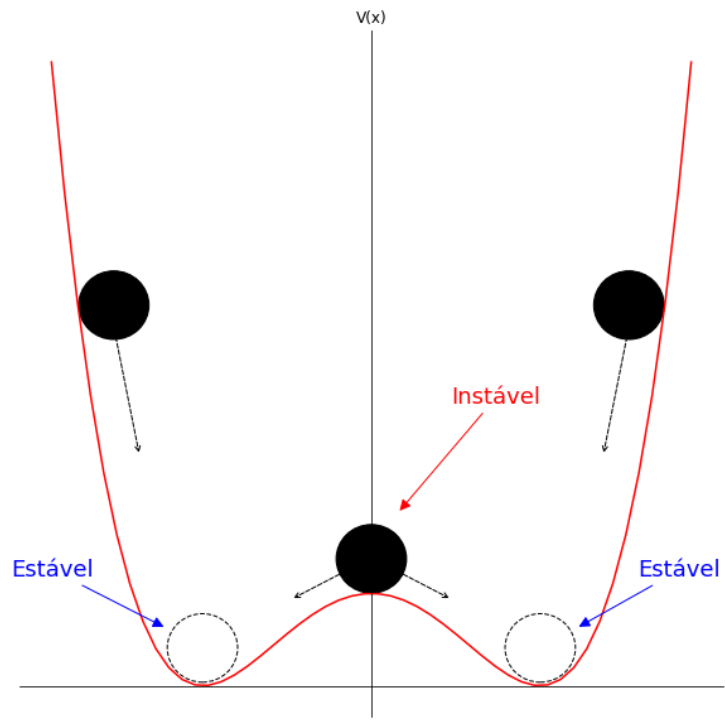


Figura 2 – Função potencial $V(x) = x^4 - x^2$, com dois pontos de equilíbrio estáveis e um ponto de equilíbrio instável.

bifurca em dois. Se b também variar para valores negativos a bifurcação é evitada e o mínimo termina a direita. Se $a = -1$ e b variar para valores positivos a bifurcação se manifestará em um salto catastrófico para o mínimo local à esquerda.

Nessa tabela é interessante notar que caso estejamos em um sistema descrito por $a = -1$ e $b = 0$ e tivéssemos uma mudança gradual para pontos onde b vai variando para pontos negativos, chegaríamos a um determinado ponto onde, caso estivéssemos no equilíbrio da esquerda, teríamos uma mudança repentina de um equilíbrio para a direita.

O comportamento dos sistemas dinâmicos que queremos estudar são completamente determinados pelas funções potenciais, como a função $V(x)$ que foi mostrada anteriormente. Essas funções potenciais dependem de variáveis comportamentais e de controle. As variáveis comportamentais descrevem o estado do sistema (variáveis de estado), enquanto as variáveis de controle determinam o comportamento do sistema. Enquanto as primeiras podem determinar se estamos em perto de um ponto de instabilidade, as últimas podem determinar para qual ponto estaremos indo.

Usaremos nesse trabalho a catástrofe do tipo *cusp*, pois é o modelo mais simples que, em termos de uma variável y_t , dá origem a descontinuidades repentinas nos estados de

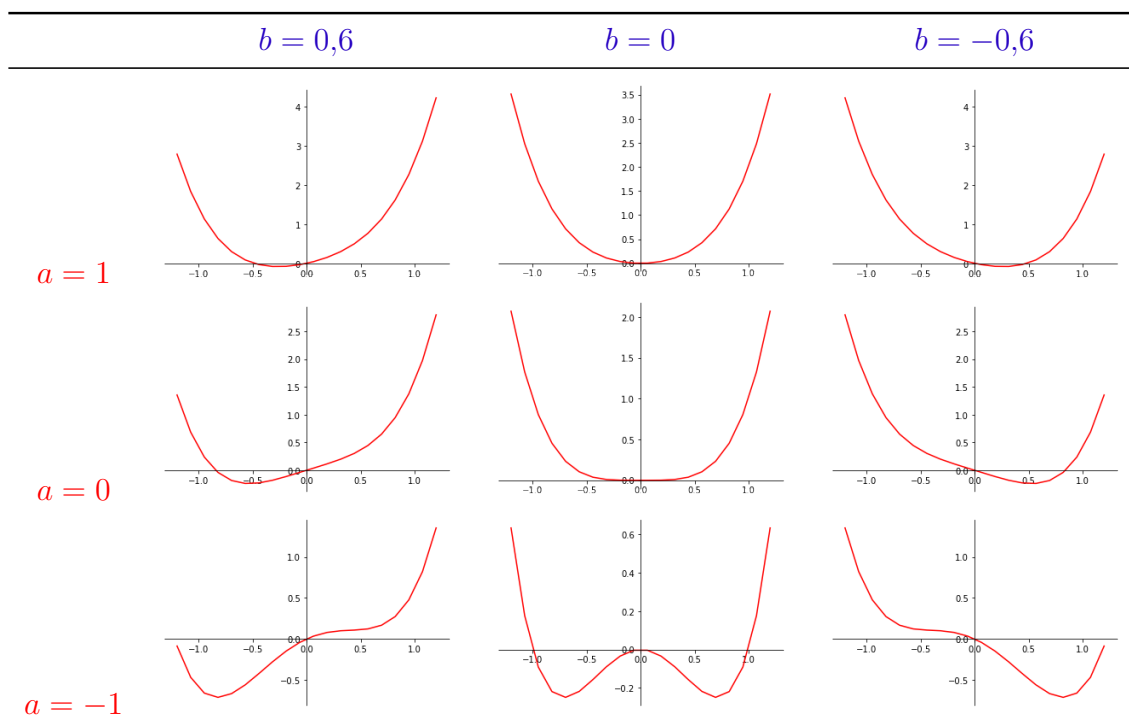


Tabela 1 – Desdobramentos do tipo *cusp* para diferentes valores de a e b .

equilíbrio dependendo de dois parâmetros de controle a e b de acordo com:

$$\frac{dy_t}{dt} = -\frac{dV(y_t; a, b)}{dy_t}, \quad (5)$$

onde $V(y_t; a, b)$ é a função potencial.

Quando a Equação 5 é zero o sistema está em equilíbrio. A superfície de equilíbrio que é formada para os parâmetros a e b foi mostrada de forma simplificada na Equação 2.

O modelo *cusp* é baseado em um sistema dinâmico determinístico, onde y_t vai variar de acordo com a derivada de função potencial *cusp* $V(y_t; a, b)$ definida por:

$$V(y_t; a, b) = -\frac{1}{4}y_t^4 + \frac{1}{2}by_t^2 + ay_t, \quad (6)$$

e no equilíbrio:

$$\frac{dV(y_t; a, b)}{dy_t} = -y_t^3 + by_t + a = 0, \quad (7)$$

onde a e b são as variáveis de controle que determinam o comportamento do sistema.

É a Equação 6 que iremos utilizar como modelo biestável para modelar os dados de criptomoedas nas seções seguintes. Esse potencial *cusp* do modelo de catástrofes é um modelo biestável que é tratado na mecânica estatística quando falamos de transições de fases, sendo um dos modelos mais simples que trazem essa biestabilidade.

As duas dimensões do espaço de controle (Figura 1), a e b , são fatores que dependerão de variáveis independentes medidas. Nesse caso, é necessário que sejam adicionadas essas variáveis independentes na análise e podemos considerar um conjunto de n variáveis

independentes $\{X_1, X_2, \dots, X_n\}$. Sendo x_i representativo da realização de X_i , as variáveis a_x e b_x são chamadas de fator de assimetria e fator de bifurcação, respectivamente:

$$a_x = a_0 + \sum_{i=1}^n a_i x_i \quad (8)$$

$$b_x = b_0 + \sum_{i=1}^n b_i x_i . \quad (9)$$

Essas equações mostram que os parâmetros a e b dependem de variáveis externas x_i . Se essas variáveis externas variarem com o tempo, os parâmetros também irão variar. Nas seções posteriores, iremos fazer essas variáveis x_i dependerem do tempo e, conseqüentemente, poderemos ter um sistema do tipo *cusp*, por exemplo, tendo seus parâmetros dependendo de variáveis que independentes que mudam com o passar do período que formos analisar.

Desta forma a Equação 7 pode ser escrita como:

$$\frac{dV(y_t; a_x, b_x)}{dy_t} = -y_t^3 + b_x y_t + a_x . \quad (10)$$

É exatamente esse tipo de potencial que iremos trabalhar, com b_x e a_x dependendo de variáveis independentes x .

Nas seções posteriores, tomaremos essas variáveis x_i como a taxa de juros do FED (x_1), valor das ações da NVIDIA (x_2) e da AMD (x_3), quantidade de dólar circulante (M2) (x_4) e o valor de fechamento do S&P 500 (x_5) ao longo de um ano; obteremos $a_0, a_1, a_2, a_3, a_4, a_5, b_0, b_1, b_2, b_3, b_4$ e b_5 ajustando $V(y_t; a, b)$ (Equação 6) à cotação diária de uma criptomoeda (y_t). Verificaremos como a e b evoluem no tempo para cada uma das criptomoedas analisadas e se os valores dessas variáveis comportamentais atingem a região de instabilidade no conjunto bifurcação (Equação 4).

3 O PROBLEMA

O problema atacado nesse trabalho de conclusão é a possível modelagem e ajuste de crises (a princípio econômicas) e os processos fora do equilíbrio que as geram. Espera-se que outros tipos de dados como sísmicos, de avalanches, vazão de rios e dados de propagação de epidemias poderão ser observados com os mesmos métodos apresentados em trabalhos futuros.

Recorre-se à um modelo de biestabilidade devido ao fato de que modelos matemáticos desse tipo podem contribuir de maneira significativa para uma boa descrição dos dados que sofrem mudanças bruscas, que é o caso das crises mencionadas. No capítulo seguinte será mostrada toda a metodologia de abordagem através do arcabouço de um modelo biestável do tipo *cusp* da teoria de catástrofes.

Na seção seguinte discutiremos o que é uma crise e quais são as crises para serem objeto de estudo desse trabalho.

3.1 O QUE É UMA CRISE FINANCEIRA

O termo crise financeira ([KINDLEBERGER, 2015](#)) é aplicado a uma variedade de situações nas quais instituições ou ativos financeiros se desvalorizam repentinamente. Algumas crises se caracterizaram pelo estouro de uma bolha financeira, outras pela quebra do mercado de ações ou por ataques especulativos à moeda, por exemplo.

Existem diversas teorias acerca do desenvolvimento das crises financeiras e como evitá-las. Entretanto, entre os economistas não é um consenso a respeito de tudo que engloba o mercado financeiro nesses períodos de crises. As crises continuam a ocorrer por todo o mundo, produzindo-se com certa regularidade. Dessa forma, pode-se pensar que elas sejam simplesmente inerentes ao funcionamento da nossa economia.

Para entendermos as crises financeiras estudadas nesse trabalho, é importante entender o que são e quais são os principais tipos de moedas que estão inseridas no nosso sistema monetário mundial.

3.2 TIPOS DE MOEDAS

Quanto aos tipos de moedas na atualidade, podemos ressaltar dois deles:

- **Moeda Mercadoria:**

É aquela que o seu valor vem da mercadoria que ela é feita. Exemplos clássicos seriam o ouro e a prata. A própria moeda, no caso, é utilizada como meio de troca.

- **Moeda de Curso Forçado:**

São aquelas moedas cuja aceitação depende da força de um decreto governamental. Esse tipo de moeda não possui valor em si mesmo, sendo seu valor apenas simbólico e sua aceitação dependente do grau de confiabilidade que as pessoas nela depositam. Exemplos são o real e o dólar.

3.3 CRIPTOMOEDAS

Uma criptomoeda é uma moeda criptografada, funcionando como um ativo digital para transações. As criptomoedas se utilizam de forte criptografia para garantir a segurança das transações financeiras que com elas ocorrem.

Em oposição às moedas digitais centralizadas e aos sistemas bancários, criptomoedas usam o controle descentralizado.

O controle descentralizado de cada criptomoeda funciona através da tecnologia de contabilidade distribuída, normalmente um *blockchain*, que serve como banco de dados de transações financeiras públicas.

3.3.1 BITCOIN

O Bitcoin (BTC), lançado pela primeira vez como software de código aberto em 2009, é geralmente considerado a primeira criptomoeda descentralizada. Desde o seu lançamento, mais de quatro mil outras criptomoedas já foram criadas.

O BTC é uma moeda que, sem ter intermediários, permite transações financeiras entre as duas partes da transação. Todas essas transações são gravadas em um banco de dados chamado *blockchain* e é muito por causa dele que a criptomoeda consegue ter um forte pilar na segurança. Outro ponto importante para a sua segurança é a presença de uma forte descentralização. Ou seja, não existe autoridade de nenhum tipo que possa manipular a emissão da moeda ou induzir algum tipo de inflação via emissão de mais dinheiro.

As oscilações sofridas pelo BTC são, como muitas moedas convencionais, devido a grandes movimentos de especulação ou variações de oferta e demanda.

3.3.2 BLOCKCHAIN

Blockchain é uma das tecnologias por trás do Bitcoin. De uma maneira bem simples, ele pode ser entendido como um grande banco de dados *online*, público e descentralizado, criado para tornar a distribuição de informação transparente e confiável, sem precisar de um agente externo e centralizador que valide esse processo. Foi graças a essas boas características que com o *blockchain* foi possível a criação de moedas digitais descentralizadas, como é o caso do Bitcoin.

O *blockchain* funciona da seguinte maneira: ele é formado por uma rede de computadores interligados (rede *peer-to-peer*). Existe um *blockchain* para cada criptomoeda (Bitcoin, Ethereum, Litecoin, etc).

No caso do Bitcoin, os computadores da rede são responsáveis por validar as transações que ocorrem, de acordo com algumas regras. Toda vez que uma transação for realizada com uma criptomoeda, ela é verificada pelos membros da rede e registrada no *blockchain*. Essas novas transações são reunidas em um bloco e criptografadas e então registradas no *blockchain*.

É dito que o *blockchain* é inviolável pois qualquer tentativa de alterar o registro de uma transação passada é facilmente identificada pelos computadores da rede.

3.3.3 MINERAÇÃO DE CRIPTOMOEDAS

Criptomoedas, diferentemente das moedas comuns, não são emitidas por bancos centrais ou governos, sendo criadas a partir de um processo chamado de mineração.

No exemplo do Bitcoin, a mineração dessa moeda é um processo finito. Isso ocorre porque o Bitcoin nasceu limitado, podendo existir 21 milhões de bitcoins no mundo. Estima-se que a última fração de bitcoin seja minerada por volta do ano de 2140.

A mineração de Bitcoin tem dois objetivos: validar as transações e registrar as transações válidas no *blockchain*. Para validar as transações, as máquinas conectadas à rede precisam resolver cálculos complexos que são ajustados pela própria rede. Somente após a solução desses cálculos, um novo bloco de transações é adicionado ao *blockchain* e o minerador recebe a recompensa. Contudo, não é simples conseguir a recompensa. Isso porque há uma grande competição entre todas as máquinas mineradoras na rede.

Nessa competição pela mineração, quem tem maior poder computacional para resolver estes cálculos leva vantagem. Além disso, a dificuldade aumenta conforme entram mais participantes na rede.

3.3.4 MERCADO DE CRIPTOMOEDAS

Na Figura 3 pode-se ver a parcela de mercado que cada um dos seis criptoativos com maior *marketshare*¹ possui dentro desse meio. O Bitcoin, no dia da consulta dos dados (01/12/2019) possuía mais de 75% do mercado global de criptomoedas.

A Figura 4 mostra a série temporal do Bitcoin desde o começo de janeiro de 2013 até setembro de 2019 obtida do site CoinMarketCap ([COINMARKETCAP, 2019](#)). Consegue-se ver alguns pontos onde o preço cai ou sobe muito rapidamente, caracterizando pontos instáveis do mercado. Esses pontos são conhecidos como pontos de altíssima volatilidade e a presença desses tipos de pontos é uma característica (não única) dos preços de criptoativos, o que torna ainda mais atrativo olhá-los pela teoria de catástrofe.

Uma questão interessante no mercado de criptomoedas é que o preço nunca para de sofrer alterações. Comparando com ativos negociados na bolsa de valores brasileira, por

¹*Marketshare* significa participação de mercado. Ou seja, é a fatia do mercado que uma empresa (ou moeda) tem no seu segmento. No caso de empresas, o *marketshare* serve para avaliar, por exemplo, a aceitação dos seus produtos.

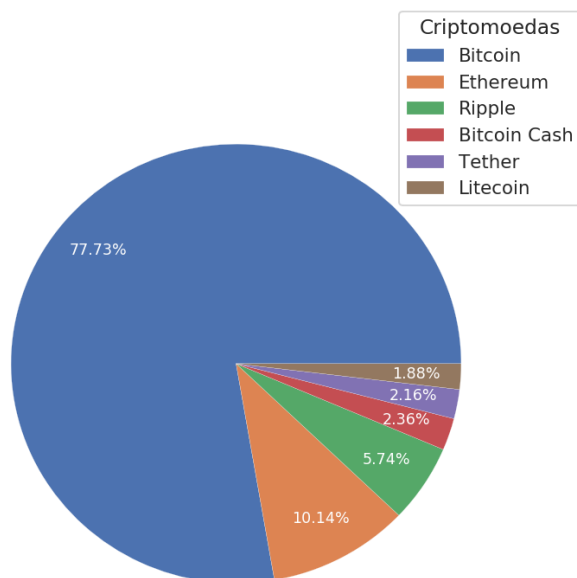


Figura 3 – Dados de *marketshare* de criptomoedas obtidas de (COINMARKETCAP, 2019).

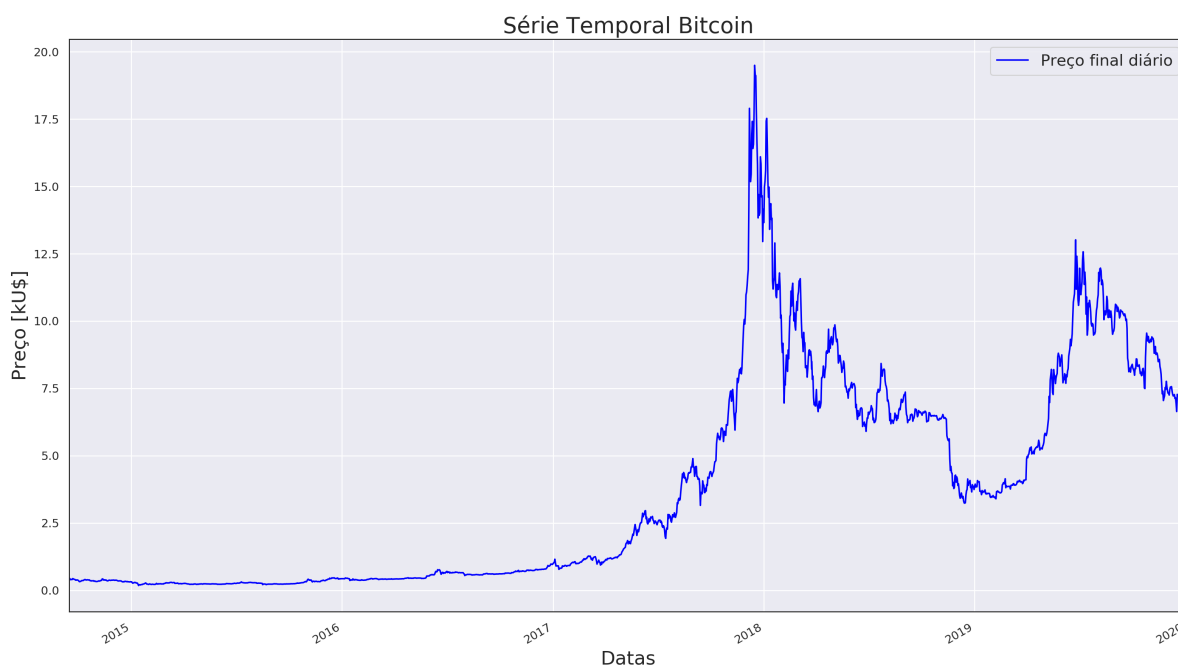


Figura 4 – Série histórica de preços de fechamento diários do Bitcoin desde 01/jan/2013 até 20/set/2019 obtidas de (COINMARKETCAP, 2019).

exemplo, que tem horário de funcionamento limitado (HORARIOS..., 2016), esses têm menos tempo diário para sofrerem alterações.

É interessante notar que a mudança absoluta no preço dos ativos não é uma quantidade

útil por si própria: uma mudança de R\$ 1,00 é bem mais significativa quando o preço do ativo é R\$ 20,00 do que quando é R\$ 200,00. Ao invés disso, a cada mudança de preço, é associado um retorno, definido como sendo a mudança no preço dividida pela sua quantidade original (WILMOTT, 1995).

No mercado financeiro, é muito comum se usar os retornos logarítmicos, pois é possível fazer a aproximação, para valores pequenos, onde o retorno é dado por uma variável r e o preço do ativo é dado por uma variável P :

$$r = \log \left(\frac{P_{t+1}}{P_t} \right) \quad (11)$$

onde \log é o logaritmo natural.

Essa aproximação pode ser feita pois como vemos na dedução a seguir, quando temos variações pequenas nos preços em relação ao intervalo de tempo o retorno percentual fica muito próximo do retorno logarítmico. Podemos fazer a aproximação em primeira ordem:

$$r = \frac{P_{t+1} - P_t}{P_t} \ll 1$$

$$e^x = \sum_{n=1}^{\infty} \frac{x^n}{n!} \approx 1 + x$$

$$x \approx \log(1 + x) \Rightarrow \log(1 + r) = \log \left(1 + \frac{P_{t+1}}{P_t} - 1 \right)$$

$$\therefore r \approx \log \left(\frac{P_{t+1}}{P_t} \right).$$

Assim, caso sejam analisados retornos, essa é a variação que utilizaremos. Essa medida relativa das mudanças é claramente um melhor indicador, e pode-se vê-la na Figura 5 para os retornos do Bitcoin no período mencionado anteriormente.

A Figura 6 mostra um histograma dos retornos logarítmicos para este ativo. Uma vantagem prática é a questão visual pois permite ver rapidamente em qual região (dos retornos) se encontra a maioria das observações (NETO, 2009). O histograma da Figura 6 sugere que a maior parte dos retornos se encontra perto do valor zero, o que já foi bem visualizado na Figura 5 ao mostrar os retornos diários do período completo.

O histograma foi feito usando um número de intervalos NI igual à diferença entre o maior e menor valor apresentado pelos retornos do bitcoin nesse período dividido pela largura h dos intervalos.

$$NI = \frac{r_{max} - r_{min}}{h}. \quad (12)$$

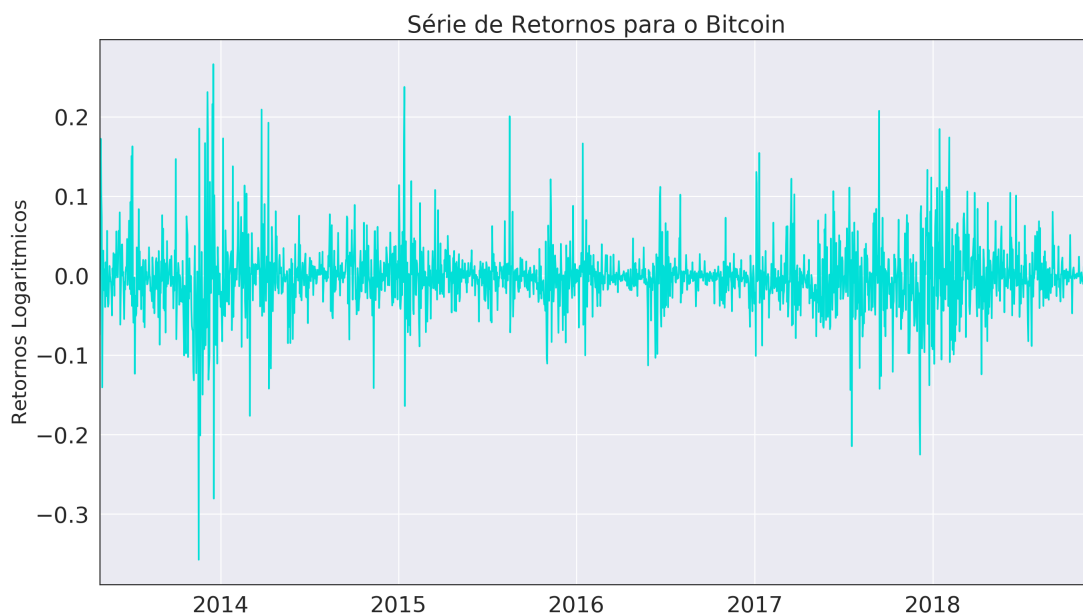


Figura 5 – Série histórica de retornos logarítmicos diários do Bitcoin desde 01/jan/2013 até 20/set/2019 obtidas de (COINMARKETCAP, 2019).

Para o número de intervalos foi utilizada a regra de Freedman-Diaconi (FREEDMAN; DIACONIS, 1981) que determina a largura dos intervalos (h) pela amplitude interquartil ($IQR(x)$) para ser usada em um histograma:

$$h = 2 \frac{IQR(x)}{n^{1/3}}. \quad (13)$$

Aqui, n representa o número de observações na amostra x que está sendo feito o histograma. IQR é uma medida de dispersão, a qual se refere a diferença entre o ponto de corte dos dados que estão entre os 25% maiores e os 25% menores da distribuição, como é possível ver na Figura 7.

Nele pode-se ver que os retornos já não parecem gaussianos. Contudo, essas afirmações não são suficientes para de fato ver que os retornos não seguem uma gaussiana. Ao decorrer do trabalho será mostrada uma forma mais concreta (qualitativa e quantitativa) de avaliar esse aspecto.

3.4 ABORDAGEM UTILIZADA PARA CRIPTOMOEDAS

Depois de um ano extremamente positivo (2017), o ano de 2018 contou com muitas quedas nos preços de várias moedas. A partir de agora analisaremos o ano de 2018 para algumas criptomoedas.

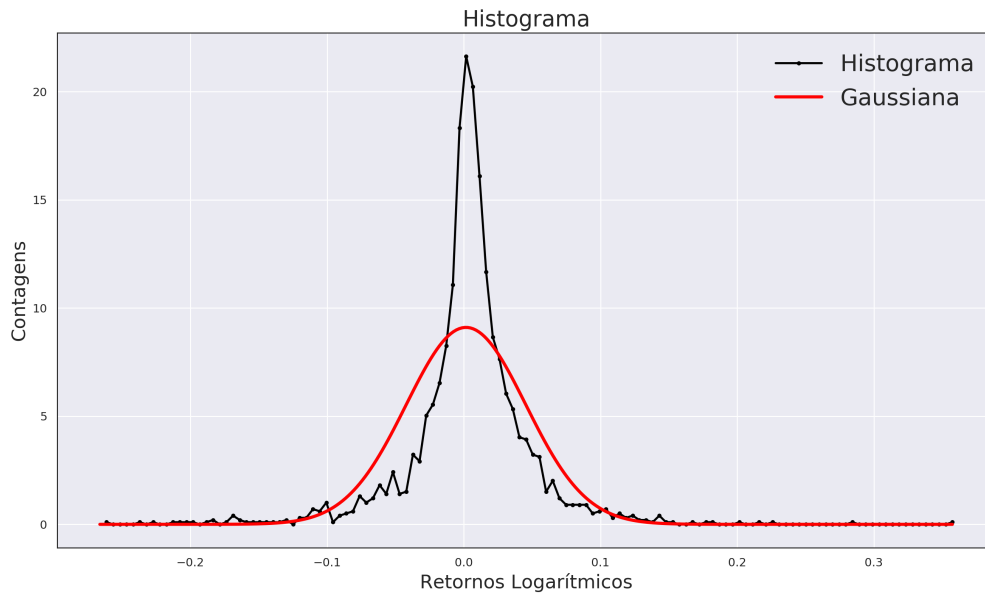


Figura 6 – Histograma dos retornos logarítmicos do Bitcoin para o período de abril de 01/jan/2013 até 20/set/2019 junto com uma função Gaussiana com a mesma média e desvio padrão. O gráfico está normalizado de forma que a área abaixo da curva do histograma seja 1.

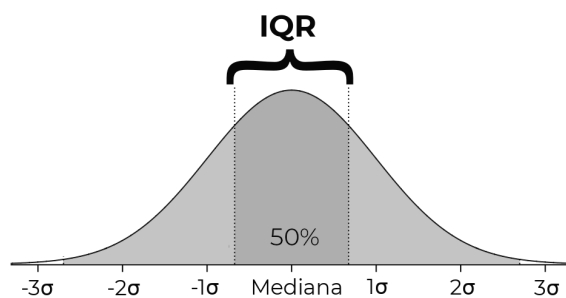


Figura 7 – Imagem representativa da distância interquartil *IQR*. Baseado em <https://en.wikipedia.org/wiki/Interquartile-range>.

Nesse trabalho, serão analisados os seguintes criptoativos no ano de 2018: Bitcoin (BTC), Ethereum (ETH), Litecoin (LTC) e Ripple (XRP). No gráfico a seguir pode-se ver a série temporal de cada uma dessas moedas ao longo de 2018. Os dados dos preços foram normalizados para que todas começassem em U\$1,00 para que pudéssemos observar todas em um único gráfico e vermos quão semelhantes elas são.

Primeiramente é interessante analisarmos algumas métricas a respeito dos dados que estamos trabalhando. O primeiro ponto de análise são os tipos de retornos que estamos

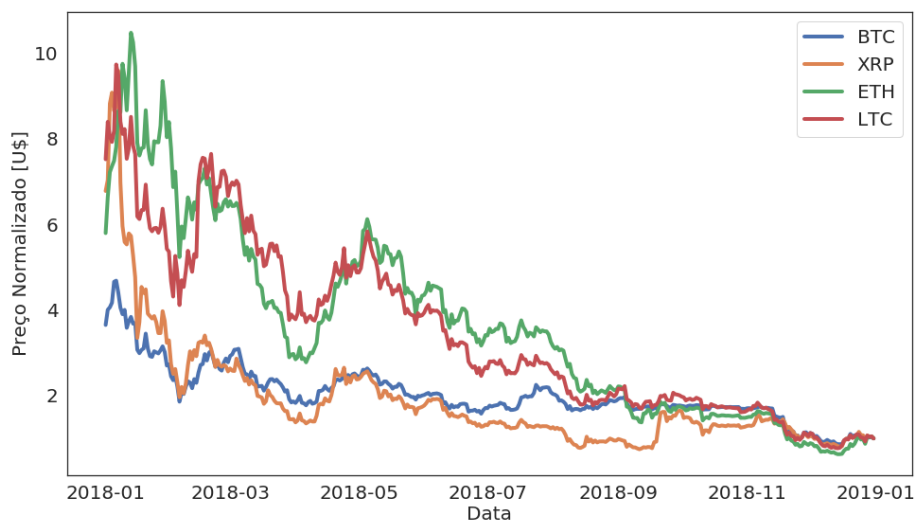


Figura 8 – Série temporal dos preços normalizados das quatro criptomoedas analisadas, de forma que todas começassem o período custando U\$ 1,00.

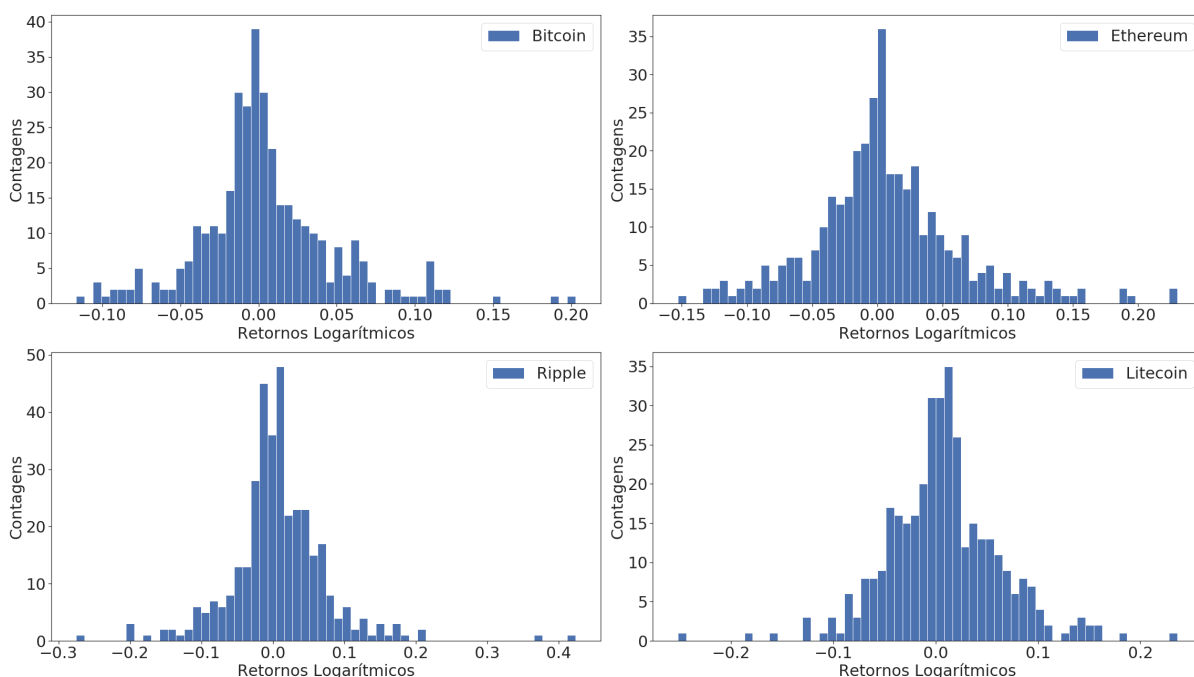


Figura 9 – Histograma de retornos das quatro criptomoedas analisadas no ano de 2018.

tratando.

A Figura 9 mostra histogramas com os mesmos parâmetros de janelas utilizados anteriormente para cada uma dessas moedas, e na Figura 10 podemos ver funções de distribuição cumulativas para esses dados comparadas a uma Gaussiana de mesma média e desvio padrão. Nessa última figura, pode-se pensar que é bem provável que não estejamos tratando de uma

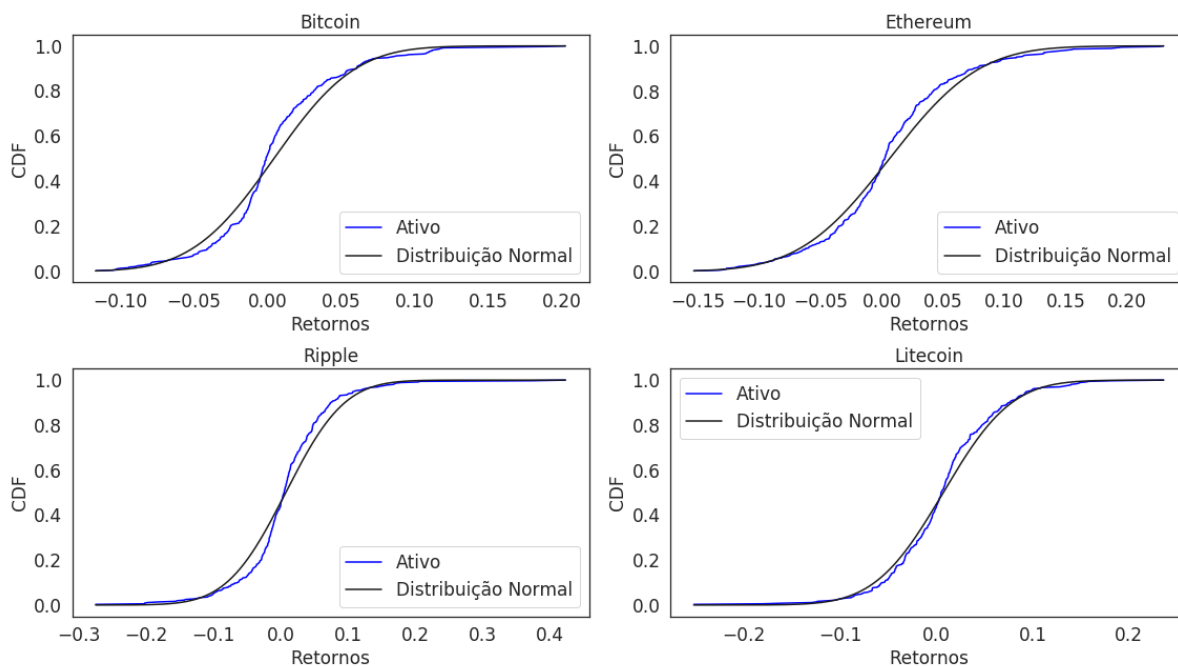


Figura 10 – Comparação da distribuição cumulativa de probabilidade para os retornos das quatro moedas e comparação com uma distribuição normal.

Tabela 2 – Tabela dos retornos ativos.

-	Média	Desvio Padrão	Curtose	Assimetria
Bitcoin	0,4491%	4,3597%	2,4845	0,7036
Ethereum	0,6455%	5,7584%	1,6706	0,5708
Litecoin	0,7105%	5,5756%	2,2502	0,0474
Ripple	0,7581%	6,8842%	6,5731	0,7343

distribuição Gaussiana, pois visualmente a distribuição cumulativa é bem diferente.

A Tabela 2 mostra algumas métricas para as distribuições de cada uma das criptomoedas e com esses valores é possível já ver algumas diferenças com relação ao que se esperaria de uma distribuição normal. Por exemplo, em uma distribuição normal (gaussiana) teríamos uma curtose igual a 3 e uma assimetria² igual a zero.

Para validação da afirmação de que as nossas distribuições de retornos não são Gaussianas pode-se fazer algumas análises estatísticas por teste de hipótese. Nesse trabalho fazemos um teste de Kolmogorov-Smirnov (uma explicação mais detalhada deste teste de hipótese pode ser visto no apêndice B).

Criptomoeda	D_n	p-value
Bitcoin	0,1206	$5,1334 \times 10^{-8}$
Ethereum	0,1531	$3,4215 \times 10^{-6}$
Ripple	0,1435	$1,4125 \times 10^{-12}$
Litecoin	0,1258	$6,4211 \times 10^{-7}$

Tabela 3 – Tabela de valores do teste de Kolgoromov-Smirnov para as criptomoedas.

3.4.1 APLICAÇÃO DO TESTE KS ÀS CRIPTOMOEDAS

Na Tabela 3 vemos essas informações³ de forma quantitativa. Visto que os valores p são muito baixos (ou D_n muito altos), podemos rejeitar a hipótese nula para todas as amostras. A tabela mostra o resultado de um teste uniamostrai, ou seja, a comparação dos dados de distribuição de retornos de cada criptomoeda com uma distribuição normal com mesma média e desvio padrão. Para todas as criptomoedas foi utilizado $n = 365$.

Com todas essas informações, é possível verificar que os retornos dos ativos a serem analisados não são oriundos de uma distribuição normal. Como discutido anteriormente, essa análise mostra que abordagens diferentes talvez sejam necessárias ao se analisar variações nos preços dessas criptomoedas. Com isso, iremos fazer uma análise levando em conta toda a discussão do capítulo anterior a respeito do modelo biestável do tipo *cusp* da teoria de catástrofes de René Thom.

²Do inglês *skewness*.

³Para o cálculo desses parâmetros foi utilizada a biblioteca *scipy* da linguagem de programação Python.

4 METODOLOGIA

Esta seção tem como objetivo explicar todos os métodos utilizados no trabalho, bem como explicar passo a passo como foi feita toda a análise do projeto.

4.1 PROJETO

Na seção 2 foi mostrada a relação dos parâmetros a e b para uma catástrofe do tipo *cusp*. Nesse tipo de sistema, os parâmetros a e b podem ser escritos, como mostramos nas Equações 8 e 9:

$$a_x = a_0 + \sum_{i=1}^n a_i x_i \quad (8)$$

$$b_x = b_0 + \sum_{i=1}^n b_i x_i . \quad (9)$$

Partindo dessas relações, deveriam ser escolhidas algumas variáveis de controle x_i para o sistema a ser estudado. Foram escolhidas algumas variáveis candidatas a serem utilizadas:

- **taxa de juros do FED (x_1):** foi utilizada a taxa de juros *overnight* americana. A taxa de juros *overnight* se refere a média das taxas aplicadas a todas as operações de crédito do dia. Os dados foram obtidos de ([Treasury Direct, 2019](#));
- **valor das ações da NVIDIA (x_2) e da AMD (x_3):** são duas empresas que fabricam placas de vídeo, *hardware* responsável pela mineração de criptomoedas. Ambas são negociadas na NASDAQ¹;
- **quantidade de dólar circulante (M2) (x_4):** é uma quantidade, em bilhões, de dólares que estão circulando no mercado. Esses dados eram de frequência semanal e foram considerados invariantes durante cada semana de análise;
- **valor de fechamento do S&P 500 (x_5):** S&P 500 é um índice americano composto por quinhentos ativos cotados nas bolsas de NYSE ou NASDAQ, qualificados devido ao seu tamanho de mercado, sua liquidez e sua representação de grupo industrial.

Dessa forma, as equações ficam:

¹É um mercado de ações automatizado norte-americano. É o segundo maior mercado de ações em capitalização de mercado do mundo, depois da Bolsa de Nova York.

$$a = a_0 + a_1x_1 + a_2x_2 + a_3x_3 + a_4x_4 + a_5x_5$$

$$b = b_0 + b_1x_1 + b_2x_2 + b_3x_3 + b_4x_4 + b_5x_5 .$$

Utilizar as cinco variáveis de parâmetros resultava em testes de hipótese refutáveis com significância menor do que utilizar alguma combinação delas. Mais a frente serão discutidas essas combinações e os resultados que elas geram, bem como uma comparação de resultados para a utilização das cinco variáveis.

4.1.1 IMPLEMENTAÇÃO DO CÓDIGO

Como um projeto que na sua visão mais ampla tem uma vertente de análise de dados, precisa-se utilizar de poder computacional para obter resultados. Ao longo do projeto foi feito um código na linguagem Python para obter, tratar, avaliar e analisar todos os dados coletados.

Foi utilizada essa linguagem devido a ser de programação livre, multiplataforma e de alto nível. Python é uma linguagem de *script* com diversas bibliotecas prontas para fazer análises numéricas, programação científica e análise de dados. É ótima para testar algoritmos e ideias antes da implementação, visto que por não ser uma linguagem compilada, ela é mais lenta para processamento.

A linguagem já foi amplamente adotada e valorizada pela comunidade científica, pois além da ampla quantidade de bibliotecas disponíveis para a análise de dados ela é muito prática e fácil de programar. A sintaxe precisa e eficiente e uma fácil curva de aprendizado fazem a linguagem produzir uma alta produtividade o que é ideal para este projeto.

4.1.1.1 AJUSTE

- **Método de estimativa da máxima verossimilhança (MLE)²**

Maximizar a função de verossimilhança determina os parâmetros que têm maior probabilidade de produzir os dados observados. De um ponto de vista estatístico, o MLE é normalmente recomendado para grandes amostras porque ele é versátil, aplicável à maioria dos modelos e diferentes tipos de dados, e produz as estimativas mais precisas.

Neste projeto, vamos utilizar o ajuste pelo método de Levenberg-Marquardt, utilizado para resolver problemas não lineares, buscando encontrar um mínimo local de uma função (uma descrição mais matemática do algoritmo de Levenberg-Marquardt encontra-se no apêndice A).

²Do inglês Maximum-Likelihood Estimation MLE.

4.2 SELEÇÃO DE VARIÁVEIS

Ao longo do desenvolvimento do trabalho observamos um melhor ajuste do modelo aos dados de cotação das criptomoedas quando foram utilizadas 4 das 5 variáveis listadas na seção 4.1. Procedemos então testando 5 combinações de 4 variáveis.

- Combinação 1: FED, NVIDIA, AMD e M2;
- Combinação 2: FED, NVIDIA, AMD, S&P500;
- Combinação 3: FED, AMD, M2, S&P500;
- Combinação 4: FED, NVIDIA, M2, S&P500;
- Combinação 5: S&P500, AMD, NVIDIA, M2.

4.2.1 DIVISÃO EM *CLUSTERS*

Outra parte interessante será o agrupamento das crises por similaridade. Por exemplo, se for possível ajustar mais de uma crise por um dos modelos de catástrofe, estes dados serão agrupados. Dentro de cada um dos grupos, tentamos novas divisões e agrupamentos de forma a ser possível ver características semelhantes entre os ativos ou dados agrupados.

Para fazer a aglomeração, utilizamos o seguinte método:

- criar vetores de a e b para cada criptomoeda;
- construir matriz de correlação;
- encontrar autovalores;
- encontrar autovetores associados.

Por exemplo, para as criptomoedas temos os seguintes vetores de a e b :

$$v_{BTC} = \begin{bmatrix} a_{BTC}(t) \\ b_{BTC}(t) \end{bmatrix}; v_{LTC} = \begin{bmatrix} a_{LTC}(t) \\ b_{LTC}(t) \end{bmatrix}; v_{XRP} = \begin{bmatrix} a_{XRP}(t) \\ b_{XRP}(t) \end{bmatrix}; v_{ETH} = \begin{bmatrix} a_{ETH}(t) \\ b_{ETH}(t) \end{bmatrix}.$$

Logo a matriz de correlação entre esses vetores torna-se:

$$\rho = \begin{bmatrix} \rho_{BTC,BTC} & \rho_{BTC,ETH} & \rho_{BTC,XRP} & \rho_{BTC,LTC} \\ \rho_{ETH,BTC} & \rho_{ETH,ETH} & \rho_{ETH,XRP} & \rho_{ETH,LTC} \\ \rho_{XRP,BTC} & \rho_{XRP,ETH} & \rho_{XRP,XRP} & \rho_{XRP,LTC} \\ \rho_{LTC,BTC} & \rho_{LTC,ETH} & \rho_{LTC,XRP} & \rho_{LTC,LTC} \end{bmatrix}$$

A correlação $\rho_{X,Y}$ do par X,Y é dada por:

$$\rho_{X,Y} = \text{corr}(X,Y) = \frac{E[(X - \mu_X)(Y - \mu_Y)]}{\sigma_X \sigma_Y},$$

onde μ_i e σ_i são a média e o desvio padrão da amostra i , respectivamente.

Tendo essa matriz de correlações entre os vetores construídos, pode-se encontrar os autovalores $\lambda_1, \lambda_2, \lambda_3, \lambda_4$ e os autovetores associados u_1, u_2, u_3 e u_4 .

Para fazer os agrupamentos, adotamos o seguinte procedimento: olhamos para os autovalores e autovetores dessa matriz de correlação ρ . Lista-se os autovetores, em ordem decrescente de autovalores associados. Do primeiro autovetor ao último, compara-se o sinal das componentes dos autovetores de forma a separar em grupos os elementos que têm componentes de sinais iguais.

Esse método para agrupamento é basicamente um método de diferença de distância que podemos ver na literatura de econofísica (MANTEGNA, 1999). Desta forma, conseguimos encontrar componentes que respondem da mesma maneira a uma perturbação externa.

4.3 ALGORITMO UTILIZADO

Para o algoritmo de Levenberg-Marquadt foi utilizado um método próprio para realizar o ajuste dos dados, inteiramente escrito em Python.

Tendo um valor prefixado para a variável λ da Equação 18, escolhemos aleatoriamente um valor para δ . Para cada iteração são feitos os cálculos do método LMA, construindo ao final de todos os passos a matriz J . Uma vez feito o ajuste, utiliza-se o resultado que ele fornece para δ para repetir o processo. O processo pode ser repetido quantas vezes for desejado ou necessário. No projeto, para cada criptomoeda foram feitas mais de 100 mil iterações.

Não necessariamente o número de vezes que se faz o ajuste deixa-o melhor, pois o resultado final é extremamente sensível aos parâmetros de entrada de δ . Por isso, muitas vezes é necessário realizar escolhas diferentes para conseguir um ajuste razoável.

4.3.1 VALIDAÇÃO DO ALGORITMO

Para validar os resultados obtidos pelo algoritmo desenvolvido em um método próprio, resolveu fazer uma verificação da capacidade de ajuste a uma função $f(x)$ bem simples. Uma validação é importante para sabermos, de fato, se o método é capaz de ajustar os dados de uma amostra com uma função qualquer desejada. Por facilidade, foi escolhido fazer o ajuste para a seguinte função:

$$f(x) = 0.5 \sin(3x) .$$

Esse problema começou juntando um vetor de 100 dados dessa função no domínio $[0,6]$, igualmente espaçados. Para acrescentar dificuldade para o algoritmo foi adicionado, a cada um desses dados, um valor de erro $\epsilon(x)$ proveniente de uma distribuição Gaussiana de média 0 e desvio padrão 1 de forma que:

$$\epsilon \sim \frac{\mathcal{N}(1,0)}{20} .$$

A divisão por 20 foi escolhida apenas para melhor caracterizar a magnitude do erro. Dessa forma, aplicando o algoritmo mencionado anteriormente, temos o resultado disposto a seguir na forma de um gráfico.

O algoritmo deveria, ao fitar os dados, conseguir convergir os parâmetros a_0 e a_1 da seguinte função para ajustar aos dados:

$$f(x) = a_0 \sin(a_1 x) .$$

O programa partiu de $a_0 = 1$ e $a_1 = 1$ pois o algoritmo de Levenberg-Marquardt precisa partir de alguma escolha inicial para os parâmetros. Em menos de 100 iterações já temos resultados consistentes, como mostrado na tabela a seguir:

	Antes do Ajuste	Depois do Ajuste
a_0	1	0,5051
a_1	1	3,0331

Tabela 4 – Tabela de valores após o ajuste de parâmetros do teste de validação realizado para o algoritmo desenvolvido.

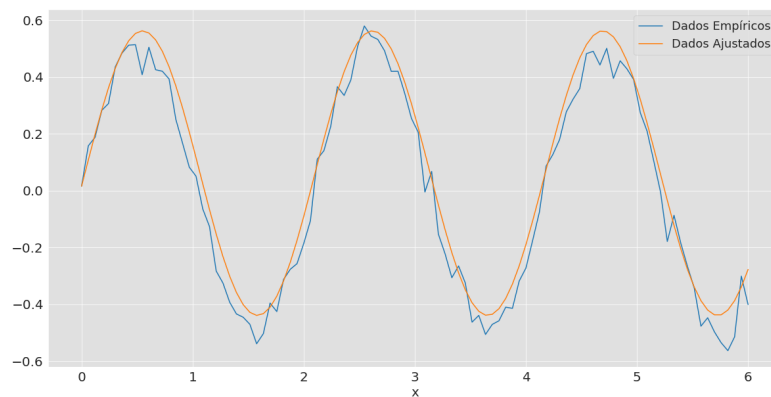


Figura 11 – Gráfico do ajuste feito para a função $f(x) = a_0 \sin(a_1 x)$ na validação do algoritmo desenvolvido.

5 ANÁLISE E DISCUSSÃO DOS RESULTADOS

Com base nos estudos apresentados na seção 2, apresentamos aqui a análise dos dados obtidos e também os resultados alcançados usando as metodologias descritas na seção anterior. Para melhor organização, esta seção será dividida em três partes: Dados, Resultados e *Clusters*.

No problema discutido anteriormente na seção 3, observou-se que os dados de variações de preços de criptoativos não seguem uma distribuição normal e, conseqüentemente, modelos que trazem abordagens diferentes podem ser úteis na hora de analisar esses ativos. A teoria de catástrofes, tratando de mudanças repentinas, pode vir a ser uma maneira interessante, como visto nas sessões anteriores, de analisar variações nos preços.

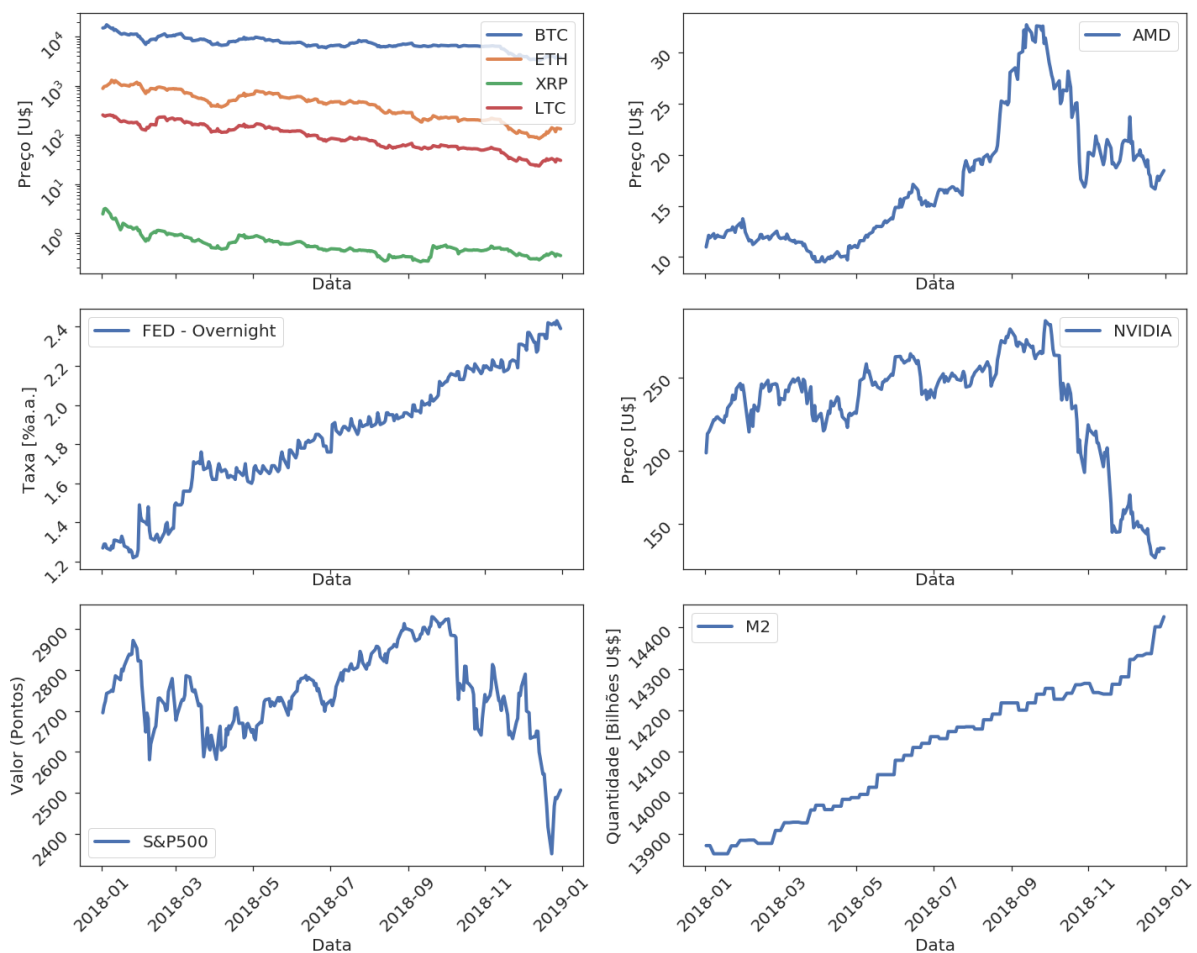


Figura 12 – Série temporal de todos os dados trabalhados nas análises.

Como dito anteriormente, todas as análises se restringem ao ano de 2018 para que se possa trazer informações mais profundas a respeito desse período. Se fossem analisados diversos anos, provavelmente qualquer análise (neste trabalho) deveria ser feita com menor profundidade. Também foram analisadas 4 criptomoedas: Bitcoin (BTC), Ripple (XRP), Litecoin (LTC) e

Ethereum (ETH), escolhidas devido a estarem entre as principais criptomoedas em questão de volume de mercado.

5.1 DADOS

Para conhecimento dos dados que estão sendo mencionados, a Figura 12 mostra um resumo da série temporal de todos os dados trabalhados. O primeiro gráfico está em escala logarítmica para facilidade de compreensão, visto que o preço nominal dos ativos é extremamente diferente.

5.2 RESULTADOS OBTIDOS

Fazendo as combinações descritas na seção 4.2, foi realizado um teste KS biamostrado para verificar qual seria o ajuste que tem a distribuição mais similar a distribuição dos dados retornos logarítmicos de cada criptomoeda. Os dados podem ser vistos na Tabela 5.2, sendo realçado o ajuste considerado melhor. Nela, cada valor de D_n representa a estatística do teste KS e o valor-p é a probabilidade de significância, ou seja, a probabilidade de se conseguir uma estatística de teste igual ou superior a que foi observada na amostra. Essa última determina, em testes de hipóteses, se podemos ou não (no nível de significância escolhido) rejeitar a hipótese nula.

Modelo Cusp - Bitcoin			Modelo Cusp - Ethereum		
Combinação	D_n	valor-p	Combinação	D_n	valor-p
1	0,0876	0,2772	1	0,1195	0,0509
2	0,0836	0,3297	2	0,1075	0,1023
3	0,1314	0,0235	3	0,1513	0,0056
4	0,1513	0,0055	4	0,1155	0,0648
5	0,0936	0,1297	5	0,2669	$2,2534 \times 10^{-8}$
Todas as variáveis	0,1702	0,0018	Todas as variáveis	0,3209	$7,5162 \times 10^{-9}$

Modelo Cusp - Ripple			Modelo Cusp - Litecoin		
Combinação	D_n	valor-p	Combinação	D_n	valor-p
1	0,1314	0,0235	1	0,0717	0,5236
2	0,1235	0,0397	2	0,0677	0,5977
3	0,1035	0,1270	3	0,1155	0,0648
4	0,2828	$2,3740 \times 10^{-9}$	4	0,0756	0,4534
5	0,1872	0,0002	5	0,1141	0,0652
Todas as variáveis	0,3109	$8,1982 \times 10^{-10}$	Todas as variáveis	0,1421	0,0466

Tabela 5 – Tabela de comparação para o ajuste considerando as cinco combinações das variáveis utilizadas e a demonstração dos resultados utilizando todas as variáveis.

Os dados que têm o valor-p muito baixo podem ser considerados como não sendo da distribuição, uma vez que a hipótese nula pode ser rejeitada para altos valores de confiança. Já os dados que contêm valor-p alto, com D_n relativamente baixo, considera-se que a hipótese nula (dados serem da mesma distribuição) não pode ser rejeitada. Esse último nada diz se, de fato, estamos falando da mesma distribuição; apenas que a hipótese nula não pôde ser rejeitada.

Escolhendo-se as combinações que geraram os melhores modelos cusp para os dados de mercado, foi criada a visualização para cada um desses ajustes obtidos. Todos foram feitos com mais de 100 mil iterações no algoritmo de Levenberg-Marquardt.

Verificando os melhores ajustes aos dados selecionados, pode-se perceber que a combinação número 2 foi a mais consistente durante os processos de ajustes.

5.2.1 CLUSTERS

Utilizando-se da metodologia previamente discutida para o agrupamento de informações, podemos chegar a uma matriz com dependência temporal para a e b , respectivamente. O cálculo de correlação dessa matriz resulta em uma matriz de correlação, como é mostrado na Tabela 6.

Resolvendo o problema de autovalores dessa matriz, achamos os seguintes autovalores:

$$\lambda_1 = 3,45881872,$$

$$\lambda_2 = 0,506954062,$$

$$\lambda_3 = 3,39649550 \times 10^{-2},$$

$$\lambda_4 = 2,62265243 \times 10^{-4},$$

com os seguintes autovetores:

$$v_1 = \begin{bmatrix} 0,51183718 \\ 0,42722517 \\ 0,7211948 \\ -0,18809416 \end{bmatrix}, v_2 = \begin{bmatrix} 0,53350077 \\ -0,03725642 \\ -0,52850883 \\ -0,65929303 \end{bmatrix}, v_3 = \begin{bmatrix} 0,51928598 \\ 0,31582681 \\ -0,37275946 \\ 0,70117464 \end{bmatrix}, v_4 = \begin{bmatrix} 0,42865102 \\ -0,8463711 \\ 0,24820731 \\ 0,19572277 \end{bmatrix}.$$

	BTC	ETH	XRP	LTC
BTC	1	0,940527	0,574350	0,983172
ETH		1	0,802551	0,936615
XRP			1	0,639032
LTC				1

Tabela 6 – Matriz de correlação para as variáveis a e b .

Munido desses dados, é possível fazer a clusterização mencionada nas seções anteriores. Olhando-se para o autovetor associado ao autovalor de maior valor (λ_1), temos diferenças de sinais da segunda componente em diante para os demais autovetores.

O autovetor v_3 forma um *cluster* com v_1 e os autovetores v_2 e v_4 formam um *cluster* separado se analisadas as segundas componentes dos autovetores. Como os autovalores λ_3 e λ_4 são muito menores, o agrupamento é encerrado. Dessa forma, pode-se ver um grupo formado por Bitcoin e Ripple e outro formado por Ethereum e Litecoin. Uma forma gráfica desse agrupamento pode ser visto na Figura 13.

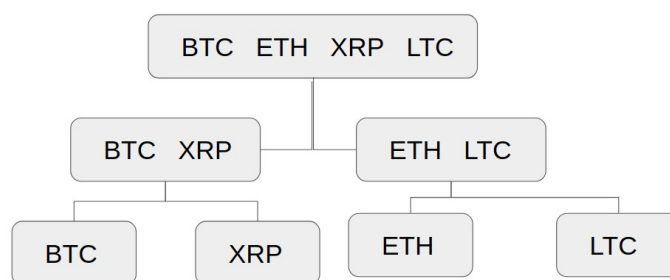


Figura 13 – *Cluster* realizado para agrupamento dos dados trabalhados.

5.2.2 RESULTADOS DOS AJUSTES

Os ajustes feitos utilizando-se dos algoritmos descritos nas sessões anteriores podem ser vistos nas Figuras 14, 15, 16 e 17. Nessas figuras é possível observar visualmente que o modelo *cusp* consegue uma distribuição cumulativa dos retornos melhor do que uma distribuição normal, corroborando com os resultados quantitativos do teste de hipótese utilizado.

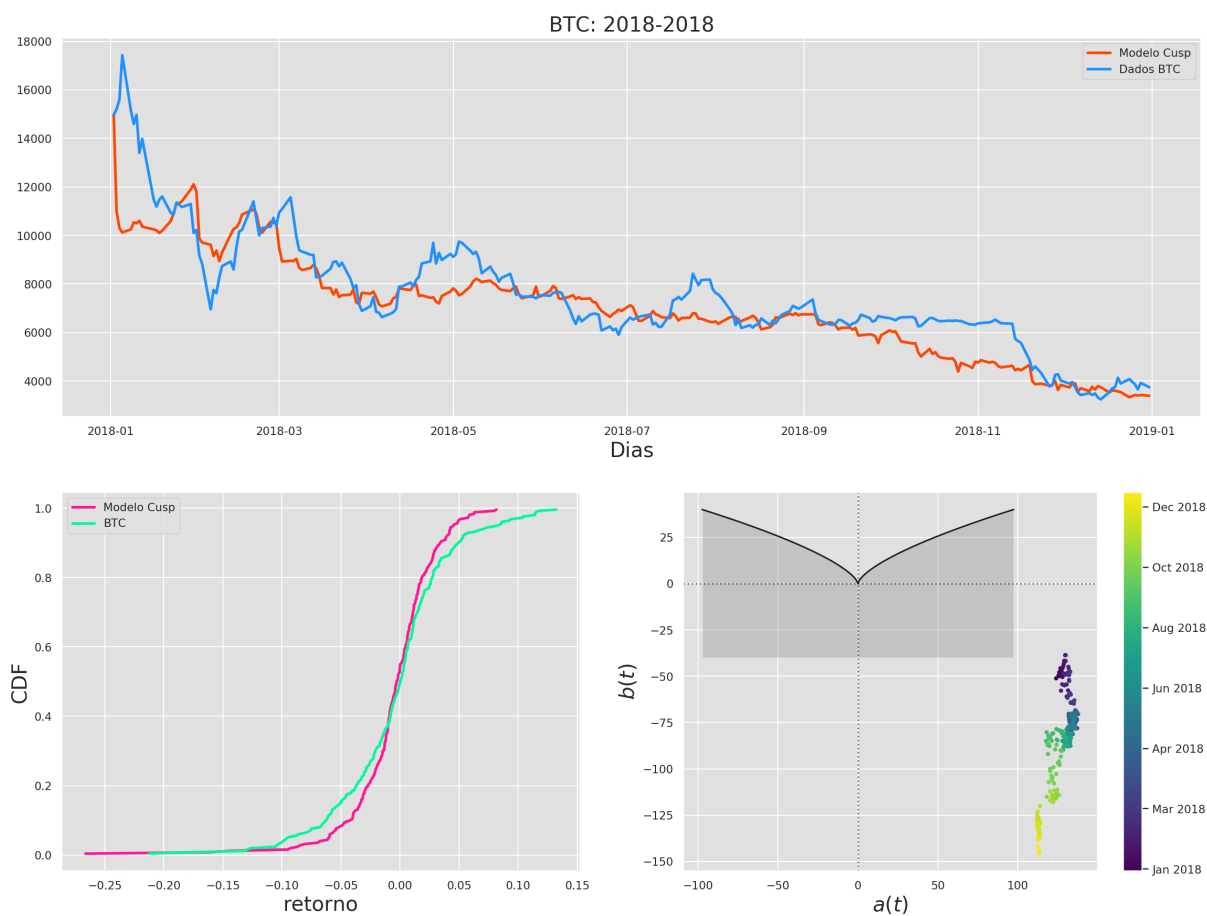


Figura 14 – Figura representando a série temporal dos dados do Bitcoin em 2018, bem como o ajuste ao método *cusp* de catástrofe. Compara-se, também, a CDF das duas distribuições e a evolução dos parâmetros a e b ao longo dos dias do ano frente à bifurcação do modelo *cusp*.

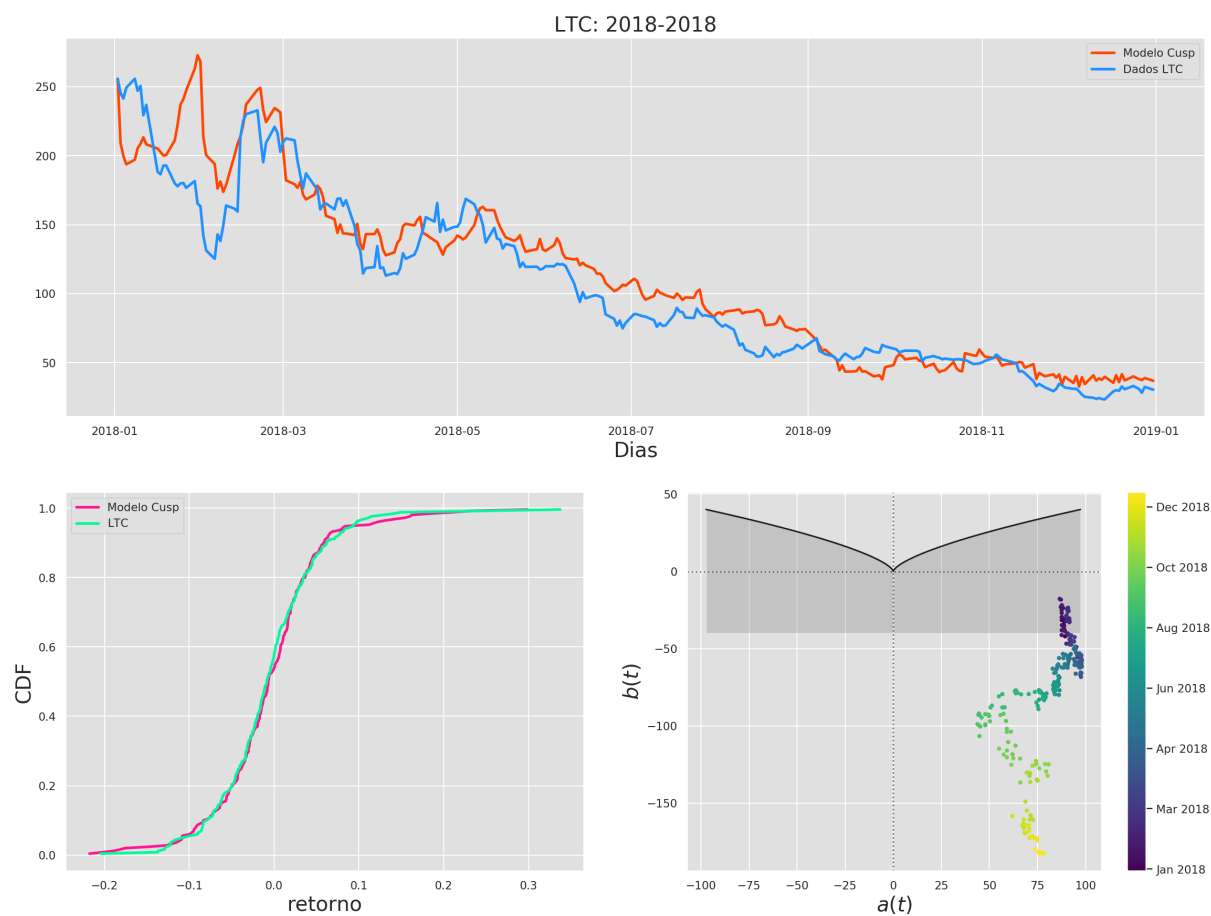


Figura 15 – Figura representando a série temporal dos dados do Litecoin em 2018, bem como o ajuste ao método *cusp* de catástrofe. Compara-se, também, a CDF das duas distribuições e a evolução dos parâmetros a e b ao longo dos dias do ano frente à bifurcação do modelo *cusp*.

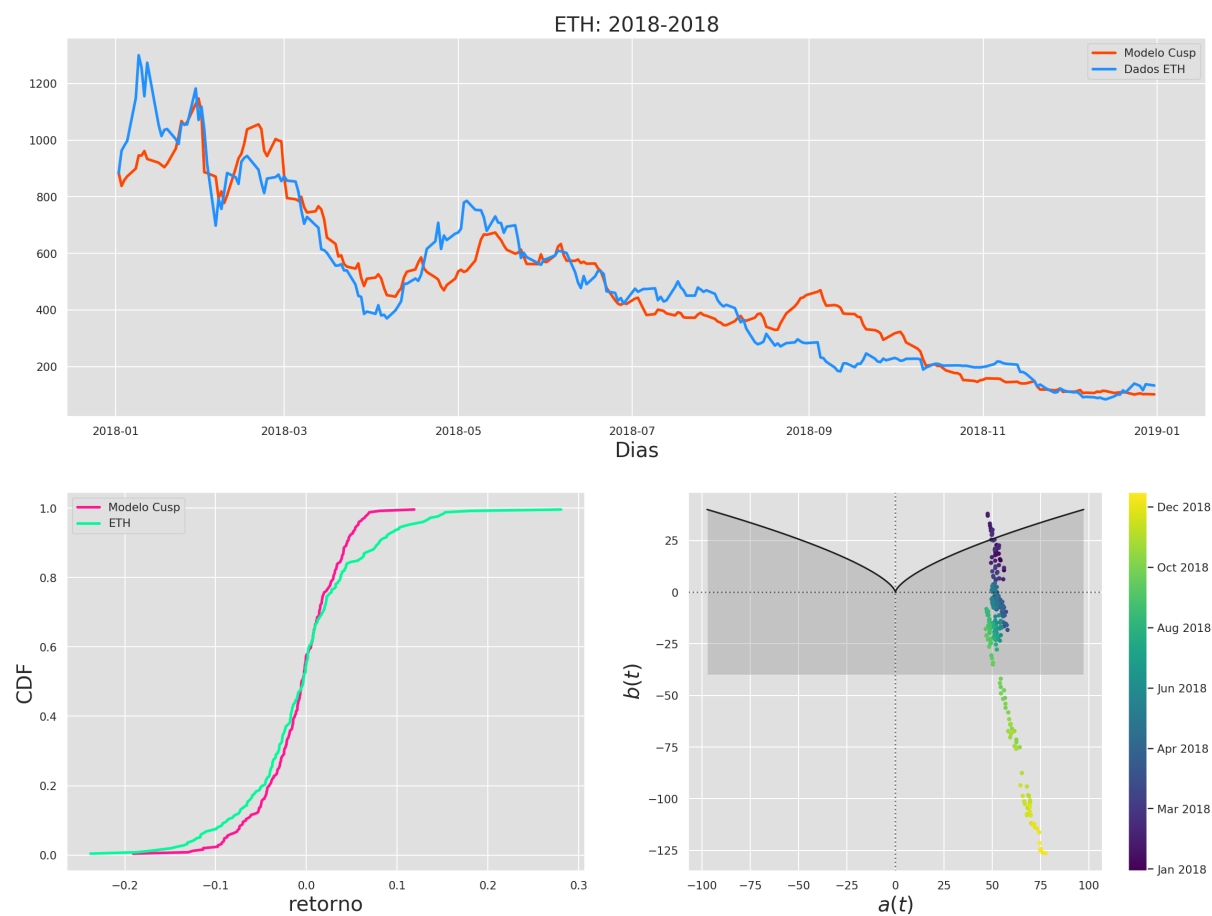


Figura 16 – Figura representando a série temporal dos dados da Ethereum em 2018, bem como o ajuste ao método *cusp* de catástrofe. Compara-se, também, a CDF das duas distribuições e a evolução dos parâmetros a e b ao longo dos dias do ano frente à bifurcação do modelo *cusp*.

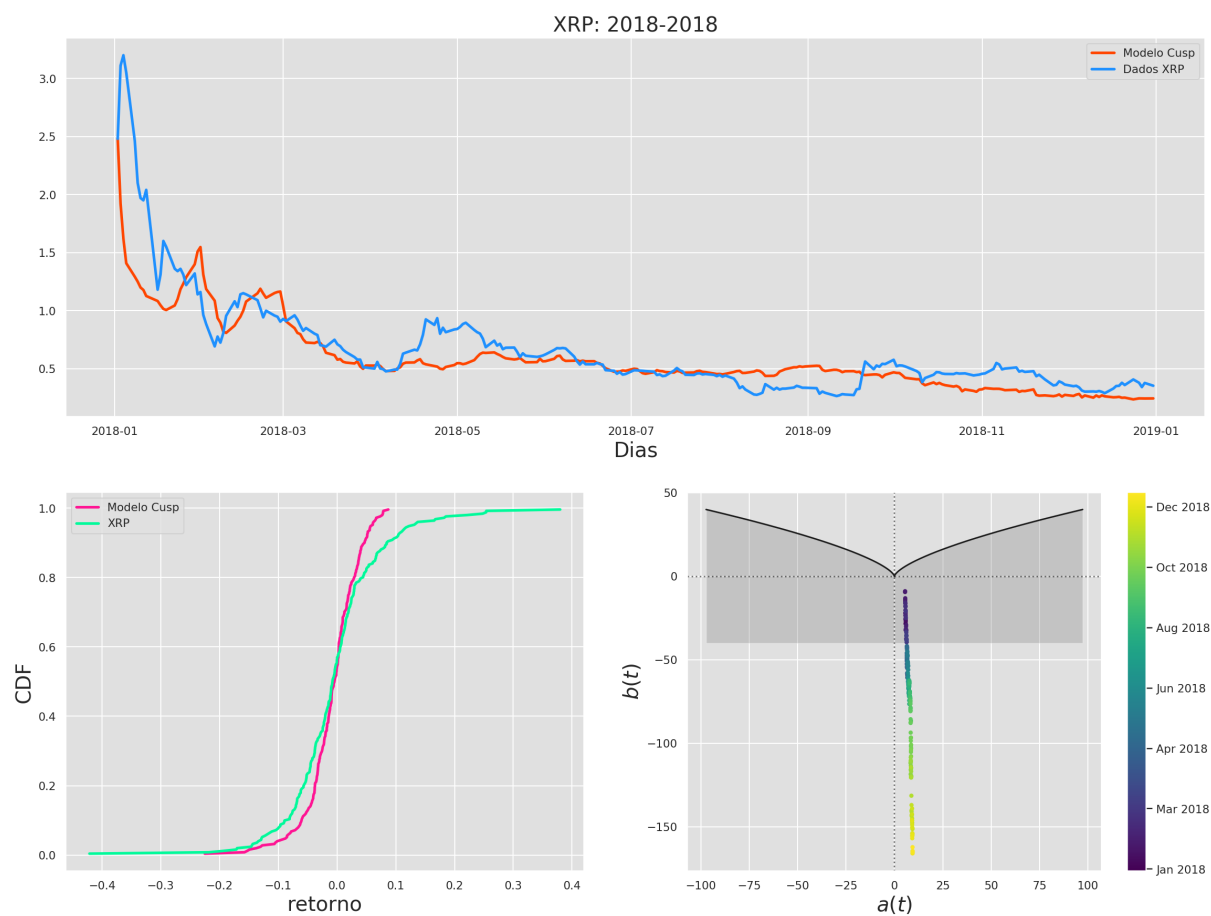


Figura 17 – Figura representando a série temporal dos dados do Ripple em 2018, bem como o ajuste ao método *cusp* de catástrofe. Compara-se, também, a CDF das duas distribuições e a evolução dos parâmetros a e b ao longo dos dias do ano frente à bifurcação do modelo *cusp*.

6 CONCLUSÃO

Observando os resultados apresentados na seção anterior, pode-se verificar que os ajustes não são perfeitos, mas conseguem dar uma boa ideia do que está acontecendo com a série temporal. Em todos os casos, vemos, de fato, um afastamento dos dados em relação a bifurcação apresentada na revisão bibliográfica. Esse afastamento quer dizer que os dados estão ficando, segundo o modelo, mais longes dos pontos instáveis do sistema. Contudo, isso não inviabiliza a utilização do modelo, pois essa informação por si só já pode ter uma relevância. Pode ser que o mercado esteve, de fato, em um momento estável.

Esse afastamento em relação a bifurcação pode ser caracterizado pela diminuição dos preços dos ativos ao longo do ano de 2018. É interessante notar a proximidade da bifurcação que os dados tinham no início do período, o que pode sugerir que a aplicação do modelo *cusp* para alguns ativos financeiros pode ser utilizada como medida de risco, da forma que a proximidade com a bifurcação pode caracterizar a proximidade do preço com um ponto instável do mercado.

6.1 TRABALHOS FUTUROS

Para trabalhos futuros, fica a possibilidade de serem avaliados períodos diferentes e períodos maiores, como por exemplo a junção dos anos de 2017 (mercado de alta) com 2018 (mercado de baixa) para criptomoedas. Outro ponto possível de melhoria seria a utilização de mais dados referentes a criptoativos, como por exemplo o volume de negociações do dia anterior em uma das variáveis de controle do modelo utilizado.

Com um trabalho levando em conta mais variáveis e mais períodos, talvez seja possível formar melhores agrupamentos e, desta forma, conseguir verificar algumas semelhanças entre os mercados de criptoativos.

A última questão que fica como uma vontade pessoal é aplicar todo esse conhecimento adquirido e também toda a parte de algoritmos desenvolvidos para olhar para mercados mais tradicionais, como o mercado de capitais, e tentar analisar de forma parecida com o que foi feito nesse trabalho de conclusão.

Referências

- ARGYRAKIS, P. What is EconoPhysics? **Journal of Engineering Science and Technology Review**, Eastern Macedonia and Thrace Institute of Technology, v. 4, n. 3, p. 207–208, dec 2011. Disponível em: <<https://doi.org/10.25103/jestr.043.01>>. Citado na página 2.
- BARUNIK, J.; VOSVRDA, M. Can a stochastic cusp catastrophe model explain stock market crashes? **Journal of Economic Dynamics and Control**, Elsevier BV, v. 33, n. 10, p. 1824–1836, oct 2009. Disponível em: <<https://doi.org/10.1016/j.jedc.2009.04.004>>. Citado na página 5.
- BELEJ, M.; KULESZA, S. Modeling the real estate prices in olsztyn under instability conditions. **Folia Oeconomica Stetinensia**, Walter de Gruyter GmbH, v. 11, n. 1, p. 61–72, jan 2012. Disponível em: <<https://doi.org/10.2478/v10031-012-0008-7>>. Citado 2 vezes nas páginas 4 e 6.
- COINMARKETCAP. 2019. Disponível em: <<https://coinmarketcap.com/>>. Acessado: 20-05-2019. Citado 3 vezes nas páginas 13, 14 e 16.
- DIKS, C.; WANG, J. Can a stochastic cusp catastrophe model explain housing market crashes? **Journal of Economic Dynamics and Control**, Elsevier BV, v. 69, p. 68–88, aug 2016. Disponível em: <<https://doi.org/10.1016/j.jedc.2016.05.008>>. Citado na página 6.
- FREEDMAN, D.; DIACONIS, P. On the histogram as a density estimator: L 2 theory. **Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete**, Citeseer, v. 57, n. 4, p. 453–476, 1981. Citado na página 16.
- HORARIOS da Bolsa. 2016. Horário da Bovespa: Confirma quando o Pregão Começa e Termina, Acessado: 10-04-2019. Disponível em: <<https://www.bussoladoinvestidor.com.br/horario-bovespa/>>. Citado na página 14.
- KINDLEBERGER, C. **Manias, Panics, and Crashes: A History of Financial Crises, Seventh Edition**. [S.l.]: Palgrave Macmillan, 2015. Citado na página 11.
- KUTNER, R. et al. Econophysics and sociophysics: Their milestones & challenges. **Physica A: Statistical Mechanics and its Applications**, Elsevier BV, v. 516, p. 240–253, feb 2019. Disponível em: <<https://doi.org/10.1016/j.physa.2018.10.019>>. Citado na página 2.
- MANTEGNA, R. N. **Introduction to Econophysics: Correlations and Complexity in Finance**. Cambridge University Press, 1999. ISBN 0521620082. Disponível em: <<https://www.xarg.org/ref/a/0521620082/>>. Citado na página 24.
- NETO, A. A. **Mercado financeiro**. [S.l.]: Atlas, 2009. ISBN 9788597001938. Citado na página 15.
- SAINT-SIMON, C.-H. **Lettres d'un habitant de Genève à ses contemporains**. Lausanne: University of Lausanne Publications, 1803. Citado na página 1.
- STEWART, I. Catastrophe theory. 1977. Citado 2 vezes nas páginas 5 e 7.
- THOM. **Stabilité structurelle et morphogénèse**. [S.l.]: Dunod, 1972. ISBN 2729600817. Citado na página 4.

Treasury Direct. **Overnight Rates**. 2019. [Online; accessed 23-June-2019]. Disponível em: <<https://www.treasurydirect.gov/GA-FI/FedInvest/selectOvernightRateDate.htm>>. Citado na página 21.

WESSELBAUM, D. Catastrophe theory and the financial crisis. **Scottish Journal of Political Economy**, Wiley, v. 64, n. 4, p. 376–391, may 2017. Disponível em: <<https://doi.org/10.1111/sjpe.12133>>. Citado na página 5.

WILMOTT, P. **The mathematics of financial derivatives : a student introduction**. Oxford New York: Cambridge University Press, 1995. ISBN 9780521497893. Citado na página 15.

ZEEMAN, E. C. Catastrophe theory. **Scientific American**, Springer Nature, v. 234, n. 4, p. 65–83, apr 1976. Disponível em: <<https://doi.org/10.1038/scientificamerican0476-65>>. Citado na página 4.

Apêndices

APÊNDICE A – MÉTODO DE LEVENBERG–MARQUARDT

O método conhecido como Método de Levenberg-Marquadt (LMA)¹ é um método utilizado para resolver problemas de mínimos quadrados não lineares, buscando encontrar um mínimo local de uma função.

Considera-se uma função de ajuste do tipo $f(\mathbf{x}; \mathbf{p})$, onde \mathbf{x} é um vetor das variáveis independentes e \mathbf{p} são os parâmetros dessa função. Faz-se, então, a seguinte aproximação:

$$f(\mathbf{x}; \mathbf{p} + \boldsymbol{\delta}) \approx f(\mathbf{x}; \mathbf{p}) + \boldsymbol{\delta} \cdot \frac{\partial f(\mathbf{x}; \mathbf{p})}{\partial \mathbf{p}} \quad (14)$$

$$\begin{bmatrix} f(x_1; [p_1 + \delta_1], [p_2 + \delta_2], \dots) \\ f(x_2; [p_1 + \delta_1], [p_2 + \delta_2], \dots) \\ \vdots \end{bmatrix} \approx \begin{bmatrix} f(x_1; p_1, p_2, \dots) \\ f(x_2; p_1, p_2, \dots) \\ \vdots \end{bmatrix} + \begin{bmatrix} \left. \frac{\partial f(x)}{\partial p_1} \right|_{x=x_1} \cdot \delta_1 + \left. \frac{\partial f(x)}{\partial p_2} \right|_{x=x_1} \cdot \delta_2 + \dots \\ \left. \frac{\partial f(x)}{\partial p_1} \right|_{x=x_2} \cdot \delta_1 + \left. \frac{\partial f(x)}{\partial p_2} \right|_{x=x_2} \cdot \delta_2 + \dots \\ \vdots \end{bmatrix}$$

$$\mathbf{F}^{n+1} \approx \begin{bmatrix} f(x_1; p_1, p_2, \dots) \\ f(x_2; p_1, p_2, \dots) \\ \vdots \end{bmatrix} + \begin{bmatrix} \left. \frac{\partial f(x)}{\partial p_1} \right|_{x=x_1} & \left. \frac{\partial f(x)}{\partial p_2} \right|_{x=x_1} & \dots \\ \left. \frac{\partial f(x)}{\partial p_1} \right|_{x=x_2} & \left. \frac{\partial f(x)}{\partial p_2} \right|_{x=x_2} & \dots \\ \vdots & \vdots & \ddots \end{bmatrix} \begin{bmatrix} \delta_1 \\ \delta_2 \\ \vdots \end{bmatrix}.$$

Ou seja, pode-se verificar essa última equação da forma:

$$\mathbf{F}^{n+1} \approx \mathbf{F}^n + \mathbf{J}^n \boldsymbol{\delta} = \mathbf{F}^n + \Delta \mathbf{F}, \quad (15)$$

onde n representa cada um dos dados que temos na amostra em questão.

Nesse caso, \mathbf{F} é um vetor coluna com os valores da função f em todos os pontos e \mathbf{J} é a matriz jacobiana para esta função. Suas colunas são as derivadas da função em relação a cada um dos parâmetros e suas linhas são estas derivadas avaliadas em cada ponto das variáveis independentes.

Para a minimização, cria-se uma função auxiliar:

$$S = |\mathbf{Y} - \mathbf{F} - \mathbf{J}\boldsymbol{\delta}|^2, \quad (16)$$

onde \mathbf{Y} é o vetor com os dados experimentais. Dessa forma, pode-se ser feito:

$$\frac{\partial S}{\partial \boldsymbol{\delta}} = 0 \quad (17)$$

$$(\mathbf{J}^T \mathbf{J})\boldsymbol{\delta} = \mathbf{J}^T(\mathbf{Y} - \mathbf{F})$$

$$(\mathbf{J}^T \mathbf{J} + \lambda \mathbf{I})\boldsymbol{\delta} = \mathbf{J}^T(\mathbf{Y} - \mathbf{F})$$

¹Do inglês *Levenberg-Marquardt Algorithm* (LMA)

$$(\mathbf{J}^T \mathbf{J} + \lambda \cdot \text{diag}(\mathbf{J}^T \mathbf{J})\delta) = \mathbf{J}^T (\mathbf{Y} - \mathbf{F}), \quad (18)$$

onde λ é o um fator de amortecimento não negativo que acaba sendo ajustado (no código) a cada iteração realizada.

APÊNDICE B – TESTE DE KOLMOGOROV-SMIRNOV

O teste de Kolmogorov-Smirnov também é conhecido como teste KS. Esse teste serve para testar a igualdade de distribuições de probabilidade e é utilizado para comparar uma amostra com uma distribuição de referência. Essa descrição se refere ao teste KS uniamostrai, contudo também existe o teste KS biamostrai. Nesse último são comparados uma amostra com a outra. O teste KS recebe esse nome em uma homenagem aos matemáticos Andrei Kolmogorov e Nikolai Smirnov.

O teste dá como resultado uma quantificação da distância entre as distribuições acumuladas em análise. Como é um teste de hipótese, gira em torno da hipótese nula (H_0) de que a amostra é retirada da distribuição de referência (no caso uniamostrai) ou de que as amostras são retiradas da mesma distribuição (no caso biamostrai).

• Teste Kolmogorov-Smirnov uniamostrai

- H_0 : A distribuição testada pode ser utilizada para prever o comportamento dos dados observados.

Para uma determinada distribuição acumulada $F(x)$, a estatística é:

$$D_n = \sup_x |F_n(x) - F(x)| . \quad (19)$$

Se a amostra vier da distribuição $F(x)$ então D_n vai convergir para zero na medida em que n for aumentando até infinito.

Nesse trabalho, foi utilizada a distribuição Gaussiana como a distribuição de referência para todos os testes uniamostrais.

• Teste Kolmogorov-Smirnov biamostrai

- H_0 : As amostras pertencem à mesma distribuição.

O teste KS também é usado para testar duas amostras diferentes. Nesse teste, faz-se a comparação para testar se duas distribuições de probabilidades (unidimensionais) são diferentes entre si. Em outras palavras, tem-se o D_n mostrado anteriormente ainda como a máxima distância entre as duas distribuições (das amostras) acumuladas, de forma que:

$$D_{n,m} = \sup |F_{1,n}(x) - F_{2,m}(x)| , \quad (20)$$

onde $F_{1,n}$ e $F_{2,m}$ são as funções distribuição da primeira e da segunda amostra e n e m os tamanhos delas, respectivamente; e \sup representa a função supremo.

A hipótese nula de que as duas amostras vêm da mesma distribuição é rejeitada ao

nível de confiança α se:

$$D_{n,m} > C(\alpha) \sqrt{\frac{n+m}{nm}} . \quad (21)$$

O valor de $C(\alpha)$ pode ser dado de forma geral por:

$$c(\alpha) = \sqrt{-\frac{1}{2} \log \left(\frac{\alpha}{2} \right)} .$$

Vale lembrar que o teste biamostrai verifica se as duas amostras de dados vêm da mesma distribuição e não especifica qual é esta distribuição comum.