UNIVERSIDADE FEDERAL DO RIO DE JANEIRO
INSTITUTO DE MATEMÁTICA
INSTITUTO TÉRCIO PACITTI DE APLICAÇÕES E PESQUISAS COMPUTACIONAIS
PROGRAMA DE PÓS-GRADUAÇÃO EM INFORMÁTICA

ALESSANDRA SENA QUADROS

# MODELOS EPIDEMIOLÓGICOS PARA PROPAGAÇÃO DE INFORMAÇÃO

Rio de Janeiro
Setembro de 2013

UNIVERSIDADE FEDERAL DO RIO DE JANEIRO
INSTITUTO DE MATEMÁTICA
INSTITUTO TÉRCIO PACITTI DE APLICAÇÕES E PESQUISAS COMPUTACIONAIS
PROGRAMA DE PÓS-GRADUAÇÃO EM INFORMÁTICA

ALESSANDRA SENA QUADROS

# MODELOS EPIDEMIOLÓGICOS PARA PROPAGAÇÃO DE INFORMAÇÃO

Dissertação de Mestrado submetida ao Corpo Docente do Departamento de Ciência da Computação do Instituto de Matemática, e Instituto Tércio Pacitti de Aplicações e Pesquisas Computacionais da Universidade Federal do Rio de Janeiro, como parte dos requisitos necessários para obtenção do título de Mestre em Informática.

Orientador: Juliana Vianna Valério
Co-orientador: Daniel Gregorio Alfaro Vigo

Rio de Janeiro
Setembro de 2013

CBIB Sena Quadros, Alessandra

Modelos Epidemiológicos para Propagação de Informação / Alessandra Sena Quadros. – Setembro de 2013.
103 f.: il.

Dissertação (Mestrado em Informática) – Universidade Federal do Rio de Janeiro, Instituto de Matemática, Instituto Tércio Pacitti de Aplicações e Pesquisas Computacionais, Programa de Pós-Graduação em Informática, Rio de Janeiro, Setembro de 2013.

Orientador: Juliana Vianna Valério.
Co-orientador: Daniel Gregorio Alfaro Vigo.

1. Modelos epidemiológicos. 2. Propagação de informação. 3. Redes sociais. – Teses. I. Vianna Valério, Juliana (Orient.). II. Gregorio Alfaro Vigo, Daniel (Co-orient.). III. Universidade Federal do Rio de Janeiro, Instituto de Matemática, Instituto Tércio Pacitti de Aplicações e Pesquisas Computacionais, Programa de Pós-Graduação em Informática. IV. Título

CDD

ALESSANDRA SENA QUADROS

# Modelos Epidemiológicos para Propagação de Informação

Dissertação de Mestrado submetida ao Corpo Docente do Departamento de Ciência da Computação do Instituto de Matemática, e Instituto Tércio Pacitti de Aplicações e Pesquisas Computacionais da Universidade Federal do Rio de Janeiro, como parte dos requisitos necessários para obtenção do título de Mestre em Informática.

Aprovado em: Rio de Janeiro, _____ de _____ de _____.

_____

Profa. Dra. Juliana Vianna Valério (Orientador)

_____

Prof. Dr. Daniel Gregorio Alfaro Vigo (Co-orientador)

_____

Prof. Dr. Mauro Antônio Rincon - PPGI/UFRJ

_____

Prof. Ph.D. Daniel Sadoc Menasché - PPGI/UFRJ

_____

Profa. Dra. Luciana Prado Mouta Pena - GMA/UFF

Rio de Janeiro
Setembro de 2013

*Dedico este trabalho às pessoas que lutam diariamente ao meu lado, transmitindo fé, amor, alegria, determinação, paciência, e coragem, tornando os meus dias mais felizes e bonitos. Aos meus pais, Reynaldo e Nora, e ao meu marido, Alex. Amo vocês!*

# AGRADECIMENTOS

Ao todo criador, Deus, que está acima de todas as coisas deste mundo. Concebendo sempre os nossos desejos e vontades, mesmo quando de forma oculta.

Aos meus pais, Reynaldo e Nora, pela confiança, amor, cuidado e sabedoria.

Ao meu amigo e esposo, Alex, por toda caminhada que fizemos juntos até o dia de hoje, e pelas próximas que virão. Pela paciência e pela compreensão, por me aturar, me ajudar e me fazer feliz.

A todos os meus amigos e colegas do mestrado, que com certeza plantaram um pedaço de si em meu coração. Mas, especialmente à Aline, pela surpresa em ser uma pessoa tão maravilhosa, ao Alan, pela paciência em seus ensinamentos e ajuda nos códigos, ao Amauri, pelos momentos de estudos no refeitório e por sempre me ajudar. Pessoas antes desconhecidas e tão diferente de mim, que me fizeram ver a vida com outros olhos, obrigada pela amizade!

Aos professores, Marcello Goulart, Luziane Ferreira, Marcos Elia, Daniel Alfaro, Adriano Joaquim, Daniel Sadoc, Jonice de Oliveira e Mauro Rincon , pela paciência, ensino, e confiança e empenho e dedicação nos ensinamentos.

Um muito obrigada a minha professora, orientadora e co-orientadora Juliana Vianna pelo aprendizado e dedicação por esses meses que trabalhamos juntas, e também pela sua confiança, compreensão e profissionalismo.

Em especial para minha querida professora Fátima Alves da graduação, que me disse um dia que somos eternos alunos.

E finalmente, agradeço a todos que me ajudaram direto ou indiretamente para o desenvolvimento deste projeto. Um MUITO OBRIGADA a todos vocês.

# RESUMO

Sena Quadros, Alessandra. **Modelos Epidemiológicos para Propagação de Informação**. Setembro de 2013. 103 f. Dissertação (Mestrado em Informática) - PPGI, Instituto de Matemática, Instituto Tércio Pacitti de Aplicações e Pesquisas Computacionais, Universidade Federal do Rio de Janeiro, Rio de Janeiro, Setembro de 2013.

Encontram-se vastamente na literatura modelos matemáticos com o objetivo de simular tanto a transmissão de doenças como a propagação de informação. Esta dissertação apresenta um modelo matemático de contágio que admite que o fenômeno de propagação de informação possa ser modelado matematicamente com os mesmos princípios que a transmissão de doenças. Baseado no modelo mais simples da epidemiologia, Suscetível - Infectado (SI), faremos uma breve análise matemática e simulações no modelo adaptado que contempla mais de uma dinâmica, e um estudo de caso para prever o valor da taxa de contágio, usando aproximação por mínimos quadrados na solução problema inverso para EDO's.

**Palavras-chave:** Modelos epidemiológicos, propagação de informação, Redes sociais.

# ABSTRACT

**Epidemiological Models for Spread Information and Social Networks**

Sena Quadros, Alessandra. **Modelos Epidemiológicos para Propagação de Informação**. Setembro de 2013. 103 f. Dissertação (Mestrado em Informática) - PPGI, Instituto de Matemática, Instituto Tércio Pacitti, Universidade Federal do Rio de Janeiro, Rio de Janeiro, Setembro de 2013.

Mathematical models are broadly used to characterize the transmission of diseases as well as the spread of information. This thesis is built on top of the classical Susceptible - Infected (SI) epidemiology model. First, we make a brief mathematical analysis of the model, establishing the stability conditions and then, we present simulations in the adapted model that includes more than one dynamic. Finally, we show a case study to predict the value of the rate of infection using the aproximation by minimum square in solving the inverse problem for ordinary differential equations.

**Keywords:** Epidemiological models, Spread information and Social Networks.

# LISTA DE FIGURAS


Figura 3.1: Fluxograma representando o modelo SI sem dinâmica vital. . . . 25
Figura 3.2: Linha de fase para $S(t)$ do modelo SI. . . . . 28
Figura 3.3: Linha de fase para $I(t)$ do modelo SI. . . . . 28
Figura 3.4: Variação temporal do número de suscetíveis S(t) e do número de infectados I(t). . . . . 30
Figura 3.5: Fluxograma representando o modelo SIR sem dinâmica vital. . . . 32
Figura 3.6: Variação temporal do número de suscetíveis $S(t)$, do número de infectados $I(t)$ e do número de recuperados $R(t)$ com $R_0 = 5,5$. . 41
Figura 3.7: Variação temporal do número de suscetíveis $S(t)$, do número de infectados $I(t)$ e do número de recuperados $R(t)$ com $R_0 = 0.33$. . 42

Figura 4.1: Fluxograma representando o Modelo ISR sem dinâmica vital. . . 46
Figura 4.2: Variação temporal do número de ignorantes $I(t)$, do número de espalhadores $S(t)$ e do número de sufocadores $R(t)$. . . . . 50
Figura 4.3: Diagrama de fluxo do Modelo de Transmissão de Informação. As setas pontilhadas indicam apenas o contato entre os indivíduos suscetíveis e infectados PACHI (2006). . . . . 52
Figura 4.4: Dinâmica do sistema. . . . . 55
Figura 4.5: Fluxograma do modelo SEIR de propagação boca-a-boca. . . . . 57
Figura 4.6: Dinâmica modelo SEIR de propagação boca-a-boca apresentado em CHA et al. (2012). . . . . 59

Figura 5.1: Fluxograma representando o modelo SI com fluxo de entrada diferente do fluxo de saída, onde apenas $S$ contribui para fluxo de entrada. . . . . 62
Figura 5.2: Variação temporal do número de suscetíveis $S(t)$ e do número de infectados $I(t)$, com $b < \mu$. . . . . 69
Figura 5.3: Distribuição da vida média do compartilhamento de links ao longo de quatro tipos diferentes de mídias sociais. (www.gestordemarkenting.com). 70
Figura 5.4: Variação temporal do número de suscetíveis $S(t)$ e do número de infectados $I(t)$, com $b > \mu$. . . . . 71
Figura 5.5: Variação temporal do número de infectados $I(t)$ após uma campanha de vacinação contra a meningite. (www.scielo.br). . . . . 71
Figura 5.6: Variação temporal do número de suscetíveis $S(t)$ e do número de infectados $I(t)$, com $b = \mu = \phi = 0$. . . . . 72
Figura 5.7: Variação temporal da tendência de busca por facebook. (www.googletrends.com acessado em 23 de julho de 2013 às 15 horas). . . . . 72

Figura 5.8: Variação temporal dos dados retirados do gráfico de tendência de busca por Facebook. . . . . 79
Figura 5.9: Variação temporal da propagação da informação com $\beta = 4,8647$ x $10^4$ estimado pelo problema inverso. . . . . 80
Figura 5.10: Gráfico de comparação - calculado pelo pelo problema inverso x dados retirados do Google Trends. . . . . 81
Figura 5.11: Gráfico de comparação com $(\beta = 4,8647$ x $10^4) * 2$. . . . . 82
Figura 5.12: Gráfico de comparação com $(\beta = 4,8647$ x $10^4)/2$. . . . . 82
Figura 5.13: Gráfico de comparação dos dados calculados e retirados da tendência de busca por Android. . . . . 84
Figura 5.14: Gráfico de comparação dos dados calculados e retirados da tendência de busca por Google. . . . . 84
Figura 5.15: Gráfico de comparação dos dados calculados e retirados da tendência de busca por Yahoo. . . . . 85
Figura 5.16: Gráfico de comparação dos dados calculados e retirados do gráfico da Figura 3.4. . . . . 86

# LISTA DE TABELAS

Tabela 5.1: : Estabilidade de cada um dos pontos de equilíbrios (5.4) e (5.5). 65

# LISTA DE ABREVIATURAS E SIGLAS

| | |
|---|---|
| SI | Suscetível - Infectado |
| SIR | Suscetível - Infectado - Recuperado |
| SEIR | Suscetível - Exposto - Infectado - Recuperado |
| DK | *Daley - Kendal* |
| ISS | *Ignorant - Spreader - Stifler* |
| PI | Problema Inverso |
| EDO's | Equações Diferenciais Ordinárias |
| RK2 | Runge-Kutta de segunda ordem |
| RK4 | Runge-Kutta de quarta ordem |

# SUMÁRIO


**1 INTRODUÇÃO** . . . . . 15

**2 EPIDEMIOLOGIA MATEMÁTICA** . . . . . 19
**2.1 Breve Histórico.** . . . . . 19

**3 MODELOS EPIDEMIOLÓGICOS CLÁSSICOS** . . . . . 22
**3.1 Modelo SI de Epidemiologia.** . . . . . 23
3.1.1 Solução Exata. . . . . 27
3.1.2 Pontos de Equilíbrio e Análise de Estabilidade. . . . . 27
3.1.3 Solução Numérica. . . . . 29
**3.2 Modelo SIR de Epidemiologia.** . . . . . 30
3.2.1 Solução Exata no Espaço de Fase. . . . . 35
3.2.2 Pontos de Equilíbrio e Análise da Estabilidade. . . . . 37
3.2.3 Solução Numérica. . . . . 40

**4 ALGUNS MODELOS MATEMÁTICOS PARA PROPAGAÇÃO DE INFORMAÇÃO** . . . . . 43
**4.1 Breve Histórico.** . . . . . 43
**4.2 Modelo ISS *Ignorant-Spreader-Stifler*.** . . . . . 45
**4.3 Modelo SIR com Dois Níveis de Atitude.** . . . . . 51
**4.4 Modelo SEIR para a propagação de informação.** . . . . . 56
**4.5 Considerações.** . . . . . 60

**5 ADAPTAÇÃO DO MODELO SUSCETíVEL - INFECTADO COM DINÂMICA VITAL PARA PROPAGAÇÃO DE INFORMAÇÃO** . . . . 61
**5.1 Pontos de Equilíbrio e Análise de Estabilidade.** . . . . . 64
**5.2 Resultados Qualitativos.** . . . . . 68
**5.3 Metodologia para o Ajuste de Parâmetros.** . . . . . 73
5.3.1 Solução do Problema Inverso. . . . . 74
**5.4 Estudo de Caso.** . . . . . 78
5.4.1 Outros Resultados. . . . . 83

**6 CONCLUSÃO** . . . . . 87

**REFERÊNCIAS** . . . . . 90

**APÊNDICE A ANÁLISE DE ESTABILIDADE DOS EQUILíBRIOS DE UMA EQUAÇÃO DIFERENCIAL ORDINÁRIA . . . . . 94**
**A.1 Linha de fase.** . . . . . 94

**APÊNDICE B ANÁLISE DE ESTABILIDADE DOS EQUILíBRIOS DE UM SISTEMA BIDIMENSIONAL . . . . . 96**
**B.1 Linearização.** . . . . . 96

**APÊNDICE C MÉTODOS DE RUNGE-KUTTA . . . . . 100**
**C.1 Método de Runge-Kutta de Segunda Ordem (RK2).** . . . . . 100
**C.2 Método de Runge-Kutta de Quarta Ordem (RK4).** . . . . . 101
C.2.1 Algoritmo - Método de Runge-Kutta de Quarta Ordem para os Sistemas de Equações Diferenciais. . . . . 101

# 1 INTRODUÇÃO

É notório o crescimento da propagação de informação com o avanço da tecnologia e uso da Internet (NAKAHARA, 2010). As mídias sociais facilitam a captação automática de dados, que podem ser usados para ajudar na construção de um modelo matemático com o objetivo de contribuir na análise da propagação da informação.

A propagação da informação ocorre em vários momentos e situações da vida. Ao ler o jornal, ouvir rádio e assistir ao noticiário, recebemos informes sobre os acontecimentos e estes são exemplos de informações transmitidas que atingem um grande número de pessoas em um curto período de tempo. Assim, podemos comparar espalhamento de ideias à deflagração da epidemia de uma doença infecciosa. Em ambos, os vírus e as notícias, são transmitidos por alguma forma de contato, e pode propagar-se de pessoa para pessoa. Esta semelhança é conhecida tanto na literatura de ciências sociais como na de epidemia (PACHI, 2006).

Há três conceitos que fazem uma ligação entre epidemias e transmissão de informação. Em primeiro lugar, o conceito de infecciosidade está presente em ambos os processos. Vírus tais como a influenza e o sarampo podem ser muito ou pouco contagiosos da mesma forma que uma notícia pode ser muito ou pouco propagada. Em segundo, ambos fazem a relação de uma pessoa de infectar a outra via contato. No caso do resfriado comum, é possível que apenas algumas tosses e espirros possam causar infecções em outra pessoa que teve contato com o vírus. O mesmo vale para a propagação de informação, a disseminação da notícia devido ao contato entre pessoas ou o contato com a notícia. A semelhança final é que os grandes eventos acontecem em um curto espaço de tempo. Uma epidemia e a informação podem atingir rapidamente a população onde elas estão inseridas.

Neste contexto social e epidemiológico, surge a motivação para este trabalho: adaptar o modelo Suscetível - Infectado (SI) a um modelo que descreva a analogia existente entre o espalhamento de ideias e as doenças infecciosas.

A finalidade inicial deste estudo foi entender os modelos usados na literatura tanto na epidemiologia quanto na propagação de informação, para fazer uma escolha do que melhor se adaptaria a uma campanha educativa, Contudo, percebemos que o modelo mais simples usado em epidemiologia, com dois compartimentos Suscetível-Infectado (SI) e com poucas constantes, pode comtemplar diferentes dinâmicas dependendo de relaçõas entre os parâmetros. Essas relações são encontradas a partir da análise de estabilidade. Lembrando que um dos desafios para usar os modelos é extrair essas constantes, que são: taxa de infecção, taxa de nascimentos, entre outras taxas que devem ser estimadas via dados retirados da informação que se deseja propagar. Sendo assim, apresentamos um maneira de aproximar as constantes do modelo a partir dos dados, utilizando o método dos mínimos quadrados na solução do problema inverso para EDO's.

Nosso objetivo é fazer uma breve análise matemática com um modelo simples que contepla mais de uma dinâmica para a propagação da informação; determinar as soluções do sistema de equações diferenciais ordinárias; analisar a dinâmica do espalhamento de ideias numa população por meio das simulações numéricas e computacionais; examinar os resultados numéricos e interpretar suas implicações com as soluções analíticas encontradas para o modelo; observar as conclusões obtidas; associar a sua utilização no modelo matemático para propagação da informação e fazer um estudo de caso utilizando o parâmetro encontrado pelo método do problema inverso para EDO's. Em resumo as principais contribuições deste trabalho são: apresentar um modelo simples com poucos parâmetros de serem estimados, porém com mais de uma dinâmica; estimar o paramêtros através de dados reais usando um dos métodos do problema inverso; aplicar o modelo em diversas siuações para

propagação de informação. Sugirimos ainda, aplicar os parâmetros das equações do modelo como *features* de um algoritmo de aprendizado por máquinas, utilizar o modelo no contexto de previsão de popularidade (SZABO; A. HUBERMAN, 2010) e de controle ótimo.

Este trabalho está organizado da seguinte forma: o capítulo 2, traz o conceito fundamental da epidemiologia e um breve histórico sobre principais modelos de epidemiologia.

No capítulo 3, cuja ênfase é dada à modelagem matemática, mostrando os modelos clássicos baseados no princípio de ação de massa, que tornou-se o mais importante conceito da epidemiologia matemática. Este conceito é traduzido pela ideia de que a disseminação da epidemia em uma população é proporcional ao produto da densidade de indivíduos suscetíveis pela densidade de indivíduos infecciosos (**?**). Os modelos Suscetível - Infectado (SI) e Suscetível - Infectado - Recuperado (SIR) usados em epidemiologia para a propagação de doenças foram analisados, pois eles assemelham-se ao problema que se deseja estudar: a propagação de uma informação. É interessante perceber que apesar da semelhança, no caso da epidemia de uma doença o objetivo é diminuir o contágio, enquanto que, se contextualizamos a propagação de informação em uma campanha educativa, o objetivo é aumentar o contágio. Ainda nesse capítulo, simulações numéricas são apresentadas.

No capítulo 4, apresenta-se alguns modelos matemáticos para propagação de informação, são eles: o modelo *Ignorant - Spreader - Stifler* (ISS), modelo SIR com dois níveis e modelo Suscetível - Exposto - Infectado - Recuperado (SEIR). Primeiramente, as características principais desses modelos são descritas e em seguida, uma reprodução dos resultados encontrados pelos autores são apresentadas.

Com o objetivo de consolidar a interdisciplinaridade entre as equações dife-

renciais, epidemiologia matemática e propagação de informação, desenvolvemos no capítulo 4 a adaptação do modelo SI para propagação da informação. Inicia-se a partir da identificação das premissas para a espalhamento das ideias e apresenta-se uma formulação do modelo e uma breve análise matemática. Simulações são feitas com o objetivo de associá-la ao estudo qualitativo do modelo adaptado e ver algumas dinâmicas para situações diferentes.

Trataremos ainda neste capítulo de uma aplicação do problema inverso, que é basicamente, um problema de estimação de parâmetros do modelo SI mais simples para a propagação de informação a partir dos dados. A dificuldade neste caso, é que o parâmetro a ser estimado é muito difícil de medir na população. Acreditamos que essa seja uma das maneiras para a solução do problema apresentado.

Ao fim deste capítulo apresentaremos um estudo de caso em uma das dinâmicas apresentadas com os efeitos da aplicação do problema inverso a dados retirados da tendências de busca por Facebook, discussão dos resultados e apresentação de outros resultados para validação do código da solução problema inverso.

Finalmente, no capítulo 5 são apresentadas as conclusões deste estudo.

# 2 EPIDEMIOLOGIA MATEMÁTICA

A palavra epidemiologia é derivada do grego (epi = sobre; demos = população, povo; logos = estudo). Desta forma, epidemiologia etimologicamente significa estudo do que ocorre em uma população. (PEREIRA, 2000).

De acordo com (MIMS, 2000), a epidemiologia é o estudo da ocorrência, disseminação e controle das doenças e tem por base a coleta de informações estatísticas detalhadas, podendo compreender desde o puramente descritivo até o analítico e experimental, nos quais os modelos matemáticos vêm assumindo cada vez mais importância. Os dados epidemiológicos podem ser usados para registrar doenças que acometem uma população e, quando infecciosas, identificar causas e formas de transmissão; e para prever a probabilidade futura da infecção, identificar fatores de risco e planejar programas de controle.

## 2.1 Breve Histórico.

Apresentaremos um breve histórico do desenvolvimento da epidemiologia matemática.

O primeiro trabalho conhecido de aplicação de matemática aos estudos da epidemiologia é atribuído a Daniel Bernoulli em 1760. Bernoulli usou um método matemático para avaliar a eficiência de políticas públicas no tratamento de varíola (BERNOULLI, 1760), (DIETZ; HEESTERBEEK, 2002).

Os cientistas Hamer e Ross, no início do século XX, investigaram a transmis-

são de doenças infecciosas por meio de simples modelos matemáticos (MASSAD, 1996). O uso de modelos matemáticos passou a ser um meio de prevenir doenças. Ross mostrou que a malária não se instala se não houver um número mínimo de mosquitos (vetores) para transmitir a doença. Em 1906, Hamer postulou que o desenvolvimento de uma epidemia depende da taxa de contato entre indivíduos não infectados (suscetíveis) e indivíduos infectados, criando assim, o modelo SI (Suscetível - Infectado). Essa noção tornou-se um dos mais importantes conceitos na aplicação de matemática em epidemiologia conhecido como princípio de ação de massa. Mais precisamente, o princípio de ação de massa que foi traduzido pela idéia de que a disseminação da epidemia em uma população é proporcional ao produto da densidade de indivíduos suscetíveis pela densidade de indivíduos infecciosos. O princípio de Hamer foi originalmente formulado através de um modelo de tempo discreto, mas, em 1908, Sir Ronald Ross (que descobriu que a malária é transmitida por mosquitos) o generalizou para tempo contínuo, em seus trabalhos sobre a dinâmica da malária (BUBNIAKOVA, 2007).

Vinte anos mais tarde, estudos matemáticos mais elaborados foram efetuados por Kermack e McKendrick, de 1927 a 1939, que desenvolveram uma teoria relacionando o surgimento de uma epidemia a um valor crítico do número de suscetíveis. O princípio de ação de massa e a teoria do valor crítico são dois marcos nos estudos da epidemiologia moderna (AMAKU, 2001), teoria esta que propôs o teorema do limiar, segundo o qual há uma densidade crítica de indivíduos suscetíveis, abaixo da qual a introdução de casos infecciosos em uma comunidade não provoca uma epidemia. O modelo que apresentou este conceito é conhecido como modelo do tipo SIR, onde S indica o compartimento dos indivíduos suscetíveis, I o grupo dos indivíduos infectados (ou também infecciosos neste caso particular) e R como o grupo dos indivíduos recuperados (ou imunes) e considera a população total constante, ou seja, a soma do número total de indivíduos distribuídos em cada um dos compartimentos é igual a população total ($S + I + R = N$).

A partir destes estudos, uma grande variedade de modelos está sendo desenvolvido e aplicado a doenças infecciosas. Os modelos matemáticos atuais envolvem aspectos como imunidade passiva, imunidade parcial, estágios de infecção, vetores de doenças, estrutura social e etária, esquemas de vacinação entre outros. Exemplos de modelos desenvolvidos para doenças como rubéola, gripe aviária, difteria, varíola, malária, dengue, herpes, sífilis, AIDS entre outras são citados nas referências deste trabalho.(KERMACK; MCKENDRICK, 1927), (ALVARENGA, 2008) e (TSUJIGUCHI, 2010).

# 3 MODELOS EPIDEMIOLÓGICOS CLÁSSICOS

Os modelos mais usados para descrever doenças transmissíveis por microparasitas (os vírus, as bactérias, os fungos e muitos protozoários) são do tipo compartimental, onde cada indivíduo de uma comunidade fechada é rotulado por seu estado de saúde em relação a alguma enfermidade. As escolhas de inclusão de compartimentos no modelo depende das características de cada doença.

Os modelos apresentados neste trabalho baseiam-se na lei de ação das massas originada do estudo de cinética química pelos noruegueses cato Guldeberg e Peter Waage em 1864, esta lei postula que "a velocidade de uma reação química é diretamente proporcional às concentrações dos reagentes". Esta lei é baseada no fato de que, numa reação química os reagentes, cada partícula de um reagente possui a mesma chance de encontro com as partículas do outro (ou outros) reagente. A interpretação desta lei para os modelos matemáticos é feita considerndo que o encontro entre variáveis é dado pelo produto delas, isto é, pelo possível número de encontro entre as variáveis de estado.

Cabe enfatizar que os modelos estudados nesta seção serão do tipo não lineares e determinístico, determinístico por que invariavelmente as mesmas entradas produzem os mesmos resultados, não contemplando a existência do acaso ou o princípio da incerteza e as variáveis são bem definidas a cada instante (RIZZI, ????).

Apresentaremos os modelos epidemiológicos clássicos que foram estudados para fazer uma melhor escolha ao modelo que se deseja adaptar para a propagação de informação, são eles: SI (Suscetível - Infectado) e SIR (Suscetível - Infectado - Recuperado), que possuem como principais hipóteses comuns aos modelos elemen-

tares:

1. A população total é considerada constante $N$. Este $N$ é suficientemente grande para que o número de indivíduos em cada categoria populacional possa ser considerado uma variável contínua à qual faça sentido aplicar taxas de variação instantâneas e onde efeitos resultantes de estocasticidade demográfica são negligíveis.

2. A população é uniforme, estando os indivíduos de cada categoria homogeneamente misturados uns com os outros; assume-se transmissão por lei da ação de massa.

3. Número reprodutivo básico, também chamado de razão básica de reprodução, denotado por $R_0$, é um número adimensional, definido como o número médio de infecções secundárias produzidas por um único indivíduo primário infectivo, em uma população inteiramente suscetível ao agente (HETHCOTE, 2000). Ele mede a velocidade inicial de crescimento da epidemia, pois cada indivíduo infectado ramifica-se em $R_0$ novos infectados que, por sua vez, originam $R_0$ infectados cada um, e assim sucessivamente.

Valores do número reprodutivo superiores a um, em um determinado grupo de risco ou população, indicam que o parasita é capaz de invadir e se estabelecer na população hospedeira. Em oposição, para valores de $R_0$ menores que um, a infecção não conseguirá se estabelecer.

## 3.1 Modelo SI de Epidemiologia.

O modelo SI (Suscetível - Infectado) é o mais simples entre os modelos de epidemia. É um modelo compartimental em que a população é dividida em apenas

dois grupos, suscetíveis ($S$) e infectados ($I$). É usado em doenças como AIDS e herpes (BRAUER; CASTILLO-CHAVEZ, 2012), (MARTIN NOWAK, 2000) e (NOWAK, 2006). Neste modelo, não há recuperado e todos na população ou são suscetíveis para a doença ou infectados. Um indivíduo infeccioso, uma vez infectado, nunca se recupera da doença.

Apresentaremos o modelo SI com população constante e admitiremos que a escala temporal da enfermidade é muito rápida quando comparada à dinâmica demográfica da população, desprezando os efeitos demográficos (nascimentos e mortes) que ocorrem dentro do período de duração da doença, o qual é normalmente dado em termos de dias ou semanas, o modelo é chamado SI sem dinâmica vital, que significa que, na escala de tempo de interesse, não ocorrem nem nascimento, nem mortes, nem qualquer tipo de migração, na população considerada.

A dinâmica do modelo SI consiste em analisar somente os indivíduos infectados, pois ele não nos fornece a recuperação dos infecciosos, assim quando um indivíduo suscetível adquire a doença, o mesmo não volta a pertencer a classe dos suscetíveis. Neste caso, também a classe de $R$ (Recuperados) está embutida em I (Infectados).

O SI é baseado nas seguintes definições:

•$S(t)$: indivíduos suscetíveis que não estão infectados, mas podem vir a estar;

•$I(t)$: indivíduos infectados que possuem o agente patogênico e podem transmití-lo a outros indivíduos;

•$N$: tamanho total da população;

•$\beta$ : taxa de contato devido o encontro com um indivíduo infectado por unidade de tempo - também chamado coeficiente de transmissão, é o número médio de contatos, de um indivíduo por unidade de tempo, em que há a transmissão da doença;

Considerando que a população total (Suscetível e infectados) está misturado homogeneamente, onde todos tem igual probabilidade de ter contato um com outro, assume-se que a taxa com a qual $S(t)$ diminui e a taxa com a qual $I(t)$ aumenta é proporcional ao produto do número de suscetíveis pelo número de infectados.

O número médio de contatos por indivíduo infeccioso por unidade de tempo ($\beta > 0$) dado o encontro com a infecção, teremos a taxa de transferência da classe S para a classe $I$, o produto $\beta SI$, esse produto vem do princípio da ação das massas.

O fluxograma para o modelo é apresentado na Figura 3.1 ao qual correspondem as seguintes equações diferenciais ordinárias:

Figura 3.1: Fluxograma representando o modelo SI sem dinâmica vital.

$$\frac{dS}{dt} = -\beta SI \tag{3.1}$$

$$\frac{dI}{dt} = \beta SI \tag{3.2}$$

Com condições iniciais: $S(0) = S_0$ e $I(0) = I_0$

Note a dimensão de $\beta$ dado que

[S] = [I] = [população] = indivíduos, que representaremos por [n]

Para obter $[\beta]$, faremos:

$$[\beta SI] = \left[\frac{dS}{dt}\right] = \left[\frac{dI}{dt}\right] = \frac{[n]}{[t]} \tag{3.3}$$

Portanto:

$[\beta] = [n]^{-1}[t]^{-1}$, onde $\beta$ é o número médio de contatos por indivíduo infeccioso por unidade de tempo.

Representando por $N = S(t) + I(t)$, somando as Equações (3.1) e (3.2), verifica-se que a população total se conserva, pois $N = constante$, então,

$$\frac{d[N]}{dt} = 0 \tag{3.4}$$

que, juntamente com a condição inicial, tem como solução:

$$N = S(0) + I(0) = S(t) + I(t) \tag{3.5}$$

### 3.1.1 Solução Exata.

De (3.5) temos, $S(t) = N - I(t)$ que substituindo em (3.2) obtemos

$$\frac{dI}{dt} = \beta I(N - I) \tag{3.6}$$

Juntamente com a condição inicial $I(0) = I_0$, fornece

$$I(t) = \frac{I_0 N}{I_0 + e^{(-N\beta t)}(N - I_0)} \tag{3.7}$$

Assim fazendo o mesmo com a Equação (3.1) encontramos como solução

$$S(t) = \frac{S_0 N}{(N - S)e^{(N\beta t)} + S_0} \tag{3.8}$$

### 3.1.2 Pontos de Equilíbrio e Análise de Estabilidade.

Fazendo $\frac{dI}{dt} = 0$ para encontrar os pontos de equilíbios $\widehat{I}$, com a condição $\left[\frac{dI}{dt}\right]_{I=\widehat{I}} = 0$, encontramos os seguintes pontos de equilíbrio:

$$\beta\widehat{I}(N - \widehat{I}) = 0 \Longleftrightarrow \widehat{I} = 0 \text{ e } \widehat{I} = N$$

O que corresponde, respectivamente $\widehat{S} = 0$ e $\widehat{S} = N$ , visto que a população total é sempre igual a $N$.

Os pontos de equilíbrios podem ser vistos na linha de fase (ver Apêndice A) que traçamos na Figura 3.2.

Figura 3.2: Linha de fase para $S(t)$ do modelo SI.

Sobre o eixo $S$, onde marcamos os pontos de equilibrio $\widehat{S} = 0$ e $\widehat{S} = N$ , temos setas apontando para direita se $\frac{dS}{dt} > 0$ e para esquerda de $\frac{dS}{dt} < 0$. A determinação do sinal de $\frac{dS}{dt}$ foi efetuada a partir de $\frac{dS}{dt} = -\beta S(N - S)$, obtida nas Equações (3.1) e (3.5).

Concluímos que $\widehat{S} = 0$, que corresponde a $\widehat{I} = N$, é um equilíbrio estável, enquanto que $\widehat{S} = N$, que corresponde a $\widehat{I} = 0$, é um ponto de equilíbrio instável do sistema.

Assim poderíamos ter partido da equação $\frac{dI}{dt} = \beta S(N - I)$, obtida a partir de (3.1) e (3.5), cuja análise de estabilidade levaria à Figura 3.3.

Figura 3.3: Linha de fase para $I(t)$ do modelo SI.

As conclusões repetem as anteriores que $\widehat{I} = N$, que corresponde a $\widehat{S} = 0$, é um equilíbrio estável, enquanto que $\widehat{I} = 0$, que corresponde a $\widehat{S} = N$, é um ponto de equilíbrio instável do sistema.

Note que quando o equilíbrio é estável, como previsto no próprio esquema compartimental, toda população ficará infectada, ou seja, haverá uma epidemia. Para o equilíbrio instável, onde, $\widehat{S} = N$ a população fica livre de doença.

### 3.1.3 Solução Numérica.

Nesta seção apresentamos uma simulação numérica que possibilita observar a dinâmica de variação temporal do modelo em uma população hipotética. O estudo da dinâmica de uma doença transmissível consiste essencialmente em esclarecer como a quantidade de indivíduos pertencentes a cada um dos compartimentos varia à medida que o tempo passa. Mesmo conhecendo a solução analítica deste sistema, resolvemos numericamente para validar o código, que poderá ser usado em sistemas cuja a solução analítica não é conhecida.

Para obter as soluções numéricas do sistema de Equações (3.1) e (3.2), utilizamos o método numérico de Runge-Kutta de quarta (Ver apêndice C) ordem cujo algoritmo foi construído no ambiente Scilab versão 5.3.3.

As condições iniciais e os parâmetros utilizados no sistema foram baseados no modelo de (BUBNIAKOVA, 2007). Assim, consideramos as condições iniciais $S_0 = 975$ e $I_0 = 25$, o parâmetro $\beta = 0.001$ e com $\triangle t = 0.1$ a solução numérica aproxima a solução analítica com a precisão da máquina. Os gráficos obtidos da simulação numérica estão representados na Figura 3.4.

De acordo com (BUBNIAKOVA, 2007) a Figura 3.4 apresenta em um mesmo sistema de eixos coordenados, os gráficos de $S(t)$ e de $I(t)$, em função do tempo $t$. Podendo observar que, em qualquer instante $t$, a soma $I(t) + S(t)$ mantém-se constante, ou seja, igual a $I_0 + S_0 = N = 1000$.

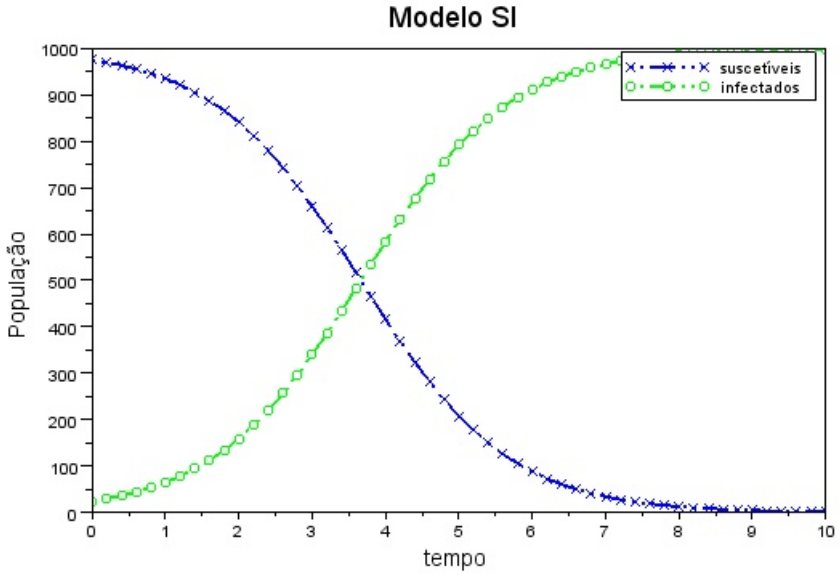


Figura 3.4: Variação temporal do número de suscetíveis S(t) e do número de infectados I(t).

## 3.2 Modelo SIR de Epidemiologia.

Os modelos matemáticos formulados por Kermack e McKendrick foram os primeiros modelos da área de epidemiologia a considerar a subdivisão da população em hospedeiros (MASSAD, 1996). Estes modelos consideram que numa população fechada, uma epidemia com microparasitas (vírus ou bactérias) ocorre através do contato direto entre indivíduos infectados e sadios. O modelo SIR a doença é transmitida pelo contato direto entre os hospedeiros que se tornam imune depois de uma simples infecção.

É usado em doenças infecciosas que ocorrem com maior frequência na infância, como rubéola, varicela, sarampo e caxumba são exemplos de doenças que costumam ser modeladas através de modelos SIR (HETHCOTE, 2000).

O modelo SIR considera uma população de tamanho constante em função do tempo. Assume-se que não existem emigração, imigração nem período latente no grupo considerado e, que a população seja homogeneamente misturada (todos tem igual probabilidade de ter contato um com o outro).

A escala temporal da enfermidade é muito pequena quando comparada à dinâmica demográfica da população. Por isso, serão desprezados os nascimentos e as mortes que ocorrem dentro do período considerado.

O modelo SIR comtempla também o compartimento R, separado de I, ou seja, permite que o indivíduo se recupere doença, ficando imune à reinfecção, este modelo nos permite computar a quantidade de infectados que se recuperam.

Ele é desenvolvido baseado nas seguintes definições:

•$S(t)$: indivíduos suscetíveis que não estão infectados, mas podem vir a estar;

•$I(t)$: indivíduos infectados que possuem o agente patogênico e podem transmití-lo a outros indivíduos;

•$R(t)$: indivíduos recuperados que se recuperaram da infecção e adquiriram imunidade, temporária ou não, à doença.

•$N$: tamanho total da população;

•$\beta$: Coeficiente de transmissão da doença, é o número médio de contatos de um indivíduo infectado numa certa unidade de tempo. É fixo e não varia sazonalmente. O tipo de contato, direto ou indireto, adequado à transmissão, depende das características específicas da doença.

•$\gamma$: Coeficiente de recuperação da doença, a variação instantânea do número de indivíduos na classe dos infecciosos ($I$) devido à recuperação da doença é proporcional ao número de infecciosos com constante de proporcionalidade $\gamma$, dada em unidades $(tempo)^{-1}$; e o período médio de infecciosidade é $\frac{1}{\gamma}$.

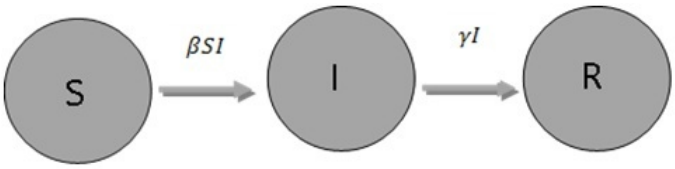


Figura 3.5: Fluxograma representando o modelo SIR sem dinâmica vital.

O fluxograma da Figura 3.5 ilustra o modelo SIR. O número de indivíduos pertencentes a cada compartimento em um dado instante $t$ é, respectivamente, $S(t)$, $I(t)$, $R(t)$, desconsiderando a heterogeneidade dos mesmos. O modelo tradicional não apresenta dinâmica vital, isto é, em qualquer instante $t$ temos que a população total $N = S(t) + I(t) + R(t)$, onde $N$ é tamanho da população, temos, que:

$$\frac{d\,[N]}{dt} = 0 \tag{3.9}$$

Desta forma, o modelo SIR é definido pelo seguinte sistema de equações diferenciais:

$$\frac{dS}{dt} = -\beta SI \tag{3.10}$$

$$\frac{dI}{dt} = \beta SI - \gamma I \tag{3.11}$$

$$\frac{dR}{dt} = -\gamma I \tag{3.12}$$

Com as condições iniciais:

$S(0) = S_0 \geqslant 0$, $I(0) = I_0 \geqslant 0$, $R(0) = R_0 = 0$

Assim, $N$ = constante, que com a condição inicial $N = S_0 + I_0$, leva a $S(t) + I(t) + R(t) = N$.

Para qualquer instante de tempo $t$, a população é conservada.

As unidades dos parâmetros envolvidos neste sistema são:

$[\beta] = [n]^{-1}\,[t]^{-1}$ e $[\gamma] = [t]^{-1}$;

quanto às variáveis dependentes temos:

$[S] = [I] = [R] = [n]$

Assim, definimos as novas variáveis dependentes adimensionais, todas variando entre 0 e 1, como:

$s \equiv \frac{S}{N}$, $i \equiv \frac{I}{N}$ e $r \equiv \frac{R}{N}$

Onde $s + i + r = 1$, e a variável independente $\tau$, através de

$\tau \equiv \beta t$

O sistema adimensionalizado é:

$$\frac{ds}{d\tau} = -\alpha si \tag{3.13}$$

$$\frac{di}{d\tau} = i(\alpha s - 1) \tag{3.14}$$

$$\frac{dr}{d\tau} = i \tag{3.15}$$

Onde definimos

$$\alpha \equiv \frac{\beta N}{\gamma} \tag{3.16}$$

Da Equação (3.11), temos que $\beta S$ é o número de contatos com transmissão da doença que resultam em aumento do número de infectados por infeccioso e por unidade de tempo; $\frac{1}{\gamma}$ é o tempo médio durante o qual um indivíduo fica infectivo, que é a capacidade de um agente de se alojar e multiplicar ou se desenvolver em um hospedeiro. Portanto, $\alpha \equiv \frac{\beta N}{\gamma}$ é o número médio de novas infecções causadas por um infectivo durante todo tempo de infecciosidade que foi encontrado no sistema adimensionalizado. Considerando uma população onde todo mundo é suscetível, então $S = N$, e considerando a $R_0$ (taxa básica de reprodução), que é definida como o

número médio de infecções secundárias produzidas quando um indivíduo infectado é introduzido em uma população inteiramente suscetível (**?**). Pode-se mostrar que um parasita é capaz de invadir e estabelecer-se em uma população de hospedeiros quando $R_0 > 1$ (epidemia acontece) e quando $R_0 < 1$ a infecção desaparece (não há epidemia). Para verificar isso, é necessário observar em que situação $\frac{di}{dt} > 0$ da Equação 3.14, ou seja, em que situação o número de indivíduos infectados aumentará. Considere a taxa básica de reprodução definida por $\alpha = R_0$. É importante destacar que mantivemos a mesma nomenclatura da taxa básica de reprodução encontrada na literatura ($R_0$), não confundir com os indivíduos recuperados ($R$).

### 3.2.1 Solução Exata no Espaço de Fase.

Neste modelo o espaço de fase é tridimensional. Analisaremos porém, apenas os subespaços SI e SR.

Ao considerarmos as trajetórias no plano de fase SI do sistema de Equações (3.10),(3.11) e (3.12) podemos obter alguns resultados analíticos.

Tomamos $I \neq 0$ e $S \neq 0$ e usamos a regra da cadeia nas duas primeiras equações do sistema de equações (3.10) e (3.11), e temos:

$$\frac{dI}{dS} = \frac{\beta S - \gamma}{-\beta S} = -1 + \frac{\gamma}{\beta S} \tag{3.17}$$

com $S(0) = S_0 > 0$ e $I(0) = I_0 > 0$.

Integramos (3.17), obtemos como solução as suas trajetórias no plano de fase SI, dadas por

$$I(S) = -S + \frac{\gamma}{\beta} \ln S + c \tag{3.18}$$

Onde $c$ é a constante de integração.

Usando as condições iniciais em $t = 0$, obtemos:

$$I(S) = -S(t) + N - \frac{\gamma}{\beta} \ln \left( \frac{S}{S_0} \right) \tag{3.19}$$

Interpretamos $I(t)$ como o número de infectados no instante $t$, que terá significado biológico se $S(t) \leqslant S_0$ e $I_0 \geqslant 0$.

Analisando a segunda equação do sistema de equações (3.11), temos

$$\left( \frac{dI}{dt} \right)_{t=0} > 0 \Longleftrightarrow I_0(\beta S_0 - \gamma) > 0 \Longleftrightarrow S_0 > \frac{\gamma}{\beta} \tag{3.20}$$

Logo, o número de infectados será crescente se a população de suscetíveis $S_0 > \frac{\gamma}{\beta}$, o que fornece $R_0 = \frac{\beta S_0}{\gamma} > 1$. O número de infectados será decrescente quando $S_0 < \frac{\gamma}{\beta}$, ou seja, $R_0 = \frac{\beta S_0}{\gamma} < 1$.

Portanto, a incidência da doença cresce se $R_0 > 1$ e decresce se $R_0 < 1$. Existe assim, um tamanho mínimo da população para que a epidemia possa ocorrer. Como o número máximo de suscetíveis é o tamanho total da população $N$, concluímos que, para ocorrer uma epidemia devemos ter $N > \frac{\gamma}{\beta}$ e, o valor $\frac{\gamma}{\beta}$ denominado limiar epidêmico, é o tamanho mínimo da população, necessário para ocorrer a epidemia.

Através do plano de fase SR, podemos determinar com varia o número de suscetíveis em função do número de removidos. Para isto, partimos das Equações (3.10) e (3.12), donde podemos (se $I \neq 0$) escrever:

$$\frac{dS}{dR} = \frac{-\alpha S}{\beta} \tag{3.21}$$

juntamente com as condições iniciais, em $t = 0$, temos: $S(R(0)) = S_0$, com $R(0) = 0$.

As soluções, de (3.21) têm a forma:

$$S = S_0 e^{\frac{-\beta}{\gamma}R} \geqslant S_0 e^{\frac{-\beta N}{\gamma}} \tag{3.22}$$

pois $N \geqslant R$.

### 3.2.2 Pontos de Equilíbrio e Análise da Estabilidade.

Os pontos de equilíbrio são do tipo $\widehat{S}$, $\widehat{I}$, $\widehat{R}$, onde $\widehat{R} = N - (\widehat{S} + \widehat{I})$ e os valores de $\widehat{S}$ e de $\widehat{I}$ satisfazem, de (3.10) e (3.11) com $\frac{dS}{dt} = 0$ e $\frac{dI}{dt} = 0$, o seguinte sistema algébrico:

$$-\beta \widehat{S}\widehat{I} = 0 \tag{3.23}$$

$$\beta \widehat{S}\widehat{I} - \gamma \widehat{I} = 0 \tag{3.24}$$

Que tem um número infinito de soluções $\widehat{I} = 0$ e $\widehat{S}$ qualquer.

Então qualquer valor de $S$, que corresponde a $I = 0$, é possível equilíbrio. Por outro lado, não há equilíbrio com número de infectados diferente de zero.

O que se pode então entender sobre a dinâmica dessas doenças?

Uma epidemia acontece quando o número de infectados cresce, ou seja, quando $\frac{dI}{dt} > 0$. Portanto, usando (3.11), temos que $\beta SI - \gamma I > 0$, o que fornece:

$$\frac{\beta S}{\gamma} > 1 \tag{3.25}$$

Considerando todo mundo suscetível, no início da epidemia, teremos para inequação (3.25), com $S = N$,

$$\frac{\beta N}{\gamma} > 1 \tag{3.26}$$

O lado esquerdo da inequação (3.26) confere com $R_0$ encontrado no sistema adimensionalizado.

A incidência da doença cresce se $R_0 > 1$ e decresce se $R_0 < 1$. Existe, portanto, um tamanho mínimo da população para que a epidemia possa ocorrer. Visto que o número de suscetíveis é o tamanho total da população $N$, concluímos que para haver uma epidemia, devemos ter $N > \frac{\gamma}{\beta}$ e, portanto, $\frac{\gamma}{\beta}$ é o tamanho mínimo da população, necessário para ocorrer a epidemia. Caso contrário não teria epidemia.

A análise de estabilidade dos pontos de equilíbrio pode ser desenvolvida através da linearização do sistema (Ver Apêndice B). As análises aqui propostas foram feitas no sistema não adimensionalizado por apresentar menos dificuldades nos cálculos.

Linearizando o sistema de Equações (3.10) e (3.11), em torno de $(\widehat{S}, \widehat{I})$, obtemos a seguinte matriz Jacobiana:

$$J(\widehat{S}, \widehat{I}) = \begin{bmatrix} -\beta\widehat{I} & -\beta\widehat{S} \\ \beta\widehat{I} & \beta\widehat{S} - \gamma \end{bmatrix} \tag{3.27}$$

Para verificar o comportamento das soluções na vizinhança do ponto de equilíbrio substituímos o ponto de equilíbrio $(\widehat{S}, \widehat{I}) = (\widehat{S}, 0)$, onde representa $S$ qualquer, em (3.27), e temos:

$$J(\widehat{S}, 0) = \begin{bmatrix} 0 & -\beta\widehat{S} \\ 0 & \beta\widehat{S} - \gamma \end{bmatrix} \tag{3.28}$$

Onde obtemos o determinante $D$, o traço $T$ e $\triangle = T^2 - 4D$, como segue:

$D = 0$, $T = \beta\widehat{S} - \gamma$ e $\triangle = (\beta\widehat{S} - \gamma)^2 > 0$

Como $D = 0$, temos um número infinito de pontos de equilíbrio (ver Apêndice B).

Analisando o sinal do traço, temos que:

$T > 0$ se $\frac{\beta\widehat{S}}{\gamma} > 1$ assim, os pontos de equilíbrio $(\widehat{S}, 0)$ serão instáveis, e:

$T < 0$ se $\frac{\beta \widehat{S}}{\gamma} < 1$, neste caso os equilíbrios serão estáveis.

A condição de estabilidade para o ponto de equilíbrio implica que a derivada de $I$, a partir de um certo $\widehat{S}$, deve ser negativa, fazendo com que a solução $I$ convirja para zero.

### 3.2.3 Solução Numérica.

A solução numérica deste modelo foi feita no software Scilab 5.3.3 e usado o Ruge-Kutta de quarta ordem como método numérico.

As condições iniciais e os parâmetros utilizados no sistema foram baseados no modelo de (TSUJIGUCHI, 2010). Assim, consideramos as condições iniciais $S_0 = 10$ e $I_0 = 1$, os parâmetros $\beta = 0.05$ e $\gamma = 0.1$ e $\triangle t = 0.1$ e apresenta erro $= 10^{-4}$, calculado pela fórmula $E =\mid max(I_1) - max(I_2) \mid$, onde $I_1$ é o máximo da função para $\triangle t = 0.1$ e $I_2$ é o máximo da função para $\triangle t = 0.1/2$. Os gráficos obtidos da simulação numérica estão representados na Figura 3.6.

Em ambas as figuras observamos que, em qualquer instante $t$, a soma $I(t) + S(t) + R(t)$ mantém-se constante. Na Figura 3.6, com $R_0 > 1$, visualiza-se que com o passar do tempo $t$, o número de suscetíveis diminui à medida de recuperados aumenta.

Na Figura 3.7, obtemos $R_0 < 1$ para as mesmas condições iniciais, $S_0 = 10$ e $I_0 = 1$ e para $\beta = 0.003$ e $\gamma = 0.1$ e $\triangle t = 0.1$ a solução numérica aproxima a solução analítica com a precisão da máquina. Observamos que com o passar do tempo $t$, o número de infetados apenas diminui (não ocorre a epidemia) até chegar a zero; por outro lado, o número de suscetíveis decresce e o número de recuperados cresce (a

população total se conserva).

É importante ressaltar que se tivermos $\gamma = 0$, retornamos ao modelo SI apresentado na seção anterior.

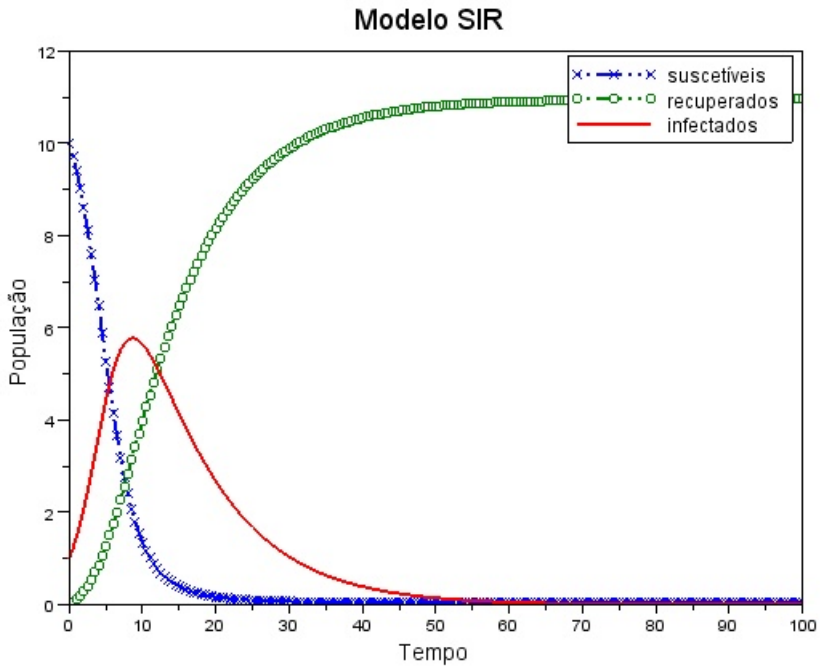


Figura 3.6: Variação temporal do número de suscetíveis $S(t)$, do número de infectados $I(t)$ e do número de recuperados $R(t)$ com $R_0 = 5,5$.

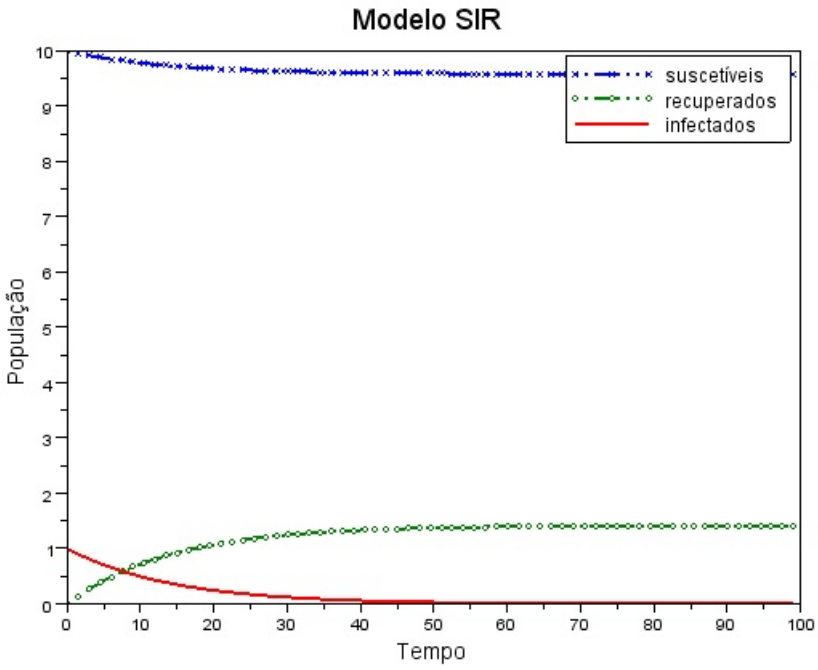


Figura 3.7: Variação temporal do número de suscetíveis $S(t)$, do número de infectados $I(t)$ e do número de recuperados $R(t)$ com $R_0 = 0.33$.

# 4 ALGUNS MODELOS MATEMÁTICOS PARA PROPAGAÇÃO DE INFORMAÇÃO

A disseminação de uma doença infecciosa ou divulgação de informações são fenômenos análogos e por isso a propagação de notícias foi modelada através da adaptação de modelos epidemiológicos.

Estudamos alguns modelos de propagação de informação para tentar entender e adaptar um modelo que possa se ajustar ao que se desejava inicialmente, um modelo para ser utilizado em uma campanha educativa.

## 4.1 Breve Histórico.

A modelagem de propagação de informação foi proposta por Rapoport em (RAPOPORT, 1953), que desenvolveu o primeiro modelo determinístico que facilitava o entendimento da comunicação em sociedade. Pitel, no trabalho (PITTEL, 1990) confirmou a conjectura do modelo Daley - Kendal, mostrando a similaridade existente na propagação de ideias e de doenças infecciosas utilizando o modelo SIR (Suscetível - Infectado - Recuperado) de epidemiologia.

Rogers, em 1995, apresentou um sistema dinâmico para modelar a propagação de ideias numa população por *memes* (E.; F.; F., 2006). Em 2002, Billings et al., (BILLINGS; SPEARS; SCHWARTZ, 2002) adaptaram dois modelos de epidemiologia em redes conectadas. A análise proporcionou uma nova visão sobre a persistência dos vírus de computador e quais as estratégias deveriam ser concebidas para o seu controle. Dodds e Watts introduziram um modelo geral de contágio, onde

sugeriu novas medidas para avaliar a suscetibilidade de uma população para eventos de grandes contágios, e também uma possível estratégia para a inibição ou para facilitar a propagação (DODDS; WATTS, 2004). Cha et al., investigaram as possíveis razões para o atraso na propagação da informação boca-a-boca e estudaram o impacto desses fatores usando o modelo SEIR(Suscetível - Exposto - Infectado - Recuperado) epidemiológico que foi adaptado para permitir a investigação da velocidade de propagação (CHA et al., 2012). Como os autores escolheram as constantes do modelo não ficou claro no artigo, mesmo com uma quantidade satisfatória de dados reais.

Muitas vezes, mesmo em posse de muitos dados, as taxas, que são fundamentais na dinâmica dos modelos, não são encontradas precisamente, comprometendo assim a eficácia desses modelos.

Pachi (PACHI, 2006) apresentou um modelo de propagação de informação com dois níveis, baseado no modelo SIR (Suscetível - Infectado - Recuperado) de epidemiologia que investiga o espalhamento da informação na população, utilizou constantes arbitrárias. Interessante observar que modelos mais complexos trazem ainda mais constantes para serem estimadas.

Os modelos matemáticos, quando bem desenvolvidos podem contribuir muito para a compreensão das possíveis dinâmicas que envolvem a propagação de informação e com isso, ser útil em diferentes situações.

A seguir apresentaremos alguns modelos matemáticos usados para a propagação de informação que foram baseados no modelo SIR de epidemiologia (seção 3.2).

## 4.2 Modelo ISS *Ignorant-Spreader-Stifler.*

O modelo ISS (*Ignorants-Spreader-Stifler*) baseado no modelo DK, e apresentado por (PIQUEIRA, 2010) e (MORENO; NEKOVEE; PACHECO, 2008), utilizou uma versão do modelo compartimental SIR de epidemiologia, onde cada um dos $N$ elementos da rede pode estar em três estados diferentes: *Ignorants* (I), *Spreader* (S) e *Stifler* (R). Onde *Ignorants* (Ignorantes) são aqueles indivíduos que não ouviram a informação e, portanto, eles são suscetíveis a serem informados. *Spreader* (Espalhadores) são os indivíduos ativos que estão espalhando a notícia. Finalmente, *Stiflers* (Sufocadores) são aqueles que conhecem o dado, mas não propagam a informação. O processo de espalhamento evolui por contatos direcionados dos espalhadores com os outros na população. Quando um espalhador encontra com um ignorante, este se transforma em um novo espalhador com probabilidade $\lambda$. A deterioração do processo de difusão pode ser devido a um mecanismo de "esquecimento" ou porque o espalhador sabe que o rumor perdeu seu "valor notícia". E o contato dos espalhadores tem probabilidade $\alpha$ de se encontrar outro espalhador ou um sufocador.

As equações do modelo são semelhantes a um modelo epidemiológico onde as populações *ignorants* e *spreader* do modelo ISS são análogos as populações suscetíveis e infectados do modelo SIR, respectivamente. Em ISS, a população Stifler são as pessoas que conhecem o boato, mas que pararam de espalhar depois de encontrar alguém que já conhece a informação, enquanto que os recuperados do modelo SIR são as pessoas que adquiriram imunidade ao vírus ou faleceram. Para evitar confusão na nomenclatura das variáveis, adotaremos (R) para Stifler, ao invés, de ($S$).

A dinâmica da transmissão da informação será então representada pelo seguinte diagrama:

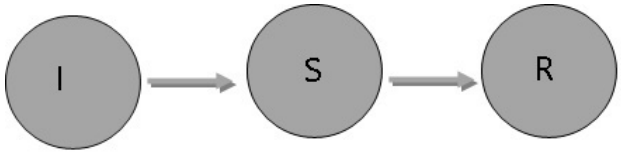


Figura 4.1: Fluxograma representando o Modelo ISR sem dinâmica vital.

O fluxograma da Figura 4.1 ilustra o modelo ISR. O número de indivíduos pertencentes a cada compartimento em um dado instante $t$ é, respectivamente, $I(t)$, $S(t)$, $R(t)$, desconsiderando a heterogeneidade dos mesmos. O modelo tradicional não apresenta dinâmica vital, isto é, em qualquer instante $t$ temos que a população total $(N)$ é constante, $I(t) + S(t) + R(t) = N$.

As relações entre os indivíduos da rede social é homogênea e para simplificar o raciocínio, $I(t) + S(t) + R(t) = N$ é considerada constante e normalizados para 1. Assim, temos que, $\frac{dN}{dt} = 0$.

Desta forma, o modelo ISR é definido pelo seguinte sistema de equações diferenciais:

$$\frac{dI}{dt} = -\beta k S I \tag{4.1}$$

$$\frac{dS}{dt} = \beta k S I - \alpha S(S + R) \tag{4.2}$$

$$\frac{dR}{dt} = \alpha S(S + R) \tag{4.3}$$

Com as condições iniciais: $I(0) = I_0 > 0$ $S(0) = S_0 > 0$ $R(0) = R_0 = 0$

Onde:

•$I(t)$ são os indivíduos *Ignorants* "não conhecem a informação";

•$S(t)$ são os indivíduos *Spreader* "espalhadores da informação";

•$R(t)$ são os indivíduos *Stifler* "sufocadores da informação";

•$\beta$ é a taxa de transmissão da informação entre os indivíduos "ignorantes" e "espalhadores";

•$\alpha$ é a taxa de recuperação ou de "esquecimento" da informação;

•$k$ é o número médio de contatos entre os indivíduos;

É importante ressaltar que, para o modelo representado pelas Equações (4.1) - (4.3), a população total da rede é $I(t)+S(t)+R(t) = N$ que permanece constante. Por conseguinte, a dimensão do espaço de estado é 2, isto é, uma das equações pode ser expressa como uma combinação linear dos outros dois.

Do sistema de Equações (4.1) - (4.3), temos que o número de espalhadores ($S$) aumenta a uma taxa proporcional a taxa de transmissão $\beta$ e o número médio de contatos $k$ de cada indivíduo. Por outro lado, o mecanismo de aniquilação considera que espalhadores decaem para a classe sufocadores ($R$) a uma taxa $\alpha k$ vezes a densidade de espalhadores e de indivíduos não-ignorantes.

Utilizando a condição de normalização, onde

$$i(t) + s(t) + r(t) = 1, \tag{4.4}$$

e as condições iniciais $i(0) = \frac{(N-1)}{N}$, $s(0) = \frac{1}{N}$ e $r(0) = 0$.

O sistema de Equações adimensionalizado será:

$$\frac{di}{dt} = -\beta k s i \tag{4.5}$$

$$\frac{ds}{dt} = \beta k s i - \alpha s(s + r) \tag{4.6}$$

$$\frac{dr}{dt} = \alpha s(s + r) \tag{4.7}$$

A dinâmica do modelo acima admite que a densidade de indivíduos espalhadores aumenta e é proporcional a taxa de espalhamento da informação $\beta$ e o número médio de contatos $k$ de cada indivíduo, para as densidades de ignorantes e indivíduos espalhadores, $i(t)$ e $s(t)$, respectivamente. Por outro lado, diminui considerando que espalhadores decaêm para a classe *stifler* a uma taxa $\alpha k$ vezes a densidade de espalhadores e de indivíduos não-ignorantes $1 - i(t) = s(t) + r(t)$.

O sistema de Equações diferenciais (4.5 - 4.7) pode ser analiticamente resolvido no tempo limite infinito quando $s(\infty) = 0$. Usando a Equação (4.4), temos que

$$\int_0^\infty s(x)\, dt = r_\infty = \lim_{x \to +\infty} r(t)$$

Introduzindo a nova variável $\beta = 1 + \frac{\lambda}{\alpha}$ nós obtemos a equação transcendental,

$$r_\infty = 1 - e^{-\beta r_\infty} \tag{4.8}$$

A Equação (4.8) admite a solução trivial $r_\infty = 0$, mas ao mesmo tempo tem uma outra solução fisicamente relevante para todos os valores dos parâmetros $\lambda$ e $\alpha$. Isto pode ser facilmente apreciada desde que a condição,

$$\frac{d}{dr_\infty}(1 - e^{-\beta r_\infty}) \mid_{r_\infty=0} > 1, \tag{4.9}$$

que se reduz para $\frac{\lambda}{\alpha} > 0$.

Este resultado, também aponta que o modelo matemático para a propagação de rumores pode ser construída de muitas maneiras diferentes.

Os pontos de equilíbrios e análise da estabilidade podem ser vistos em (PIQUEIRA, 2010).

A solução numérica com as condições iniciais e os parâmetros utilizados no sistema foram baseados em (PIQUEIRA, 2010). Assim, consideramos as condições iniciais $S_0 = 9$ e $I_0 = 1$, os parâmetros $\beta = 0.04$, $\alpha = 0.03$, $k = 0.8$ e a solução numérica se aproxima da solução analítica com a precisão da máquina, apresentando erro nulo. O gráfico obtido da simulação numérica está representado na Figura 4.2.

A análise do modelo de ISR mostra que para uma dada população $T$, os parâmetros de controle principal são as taxas $\beta$ e $\alpha$ que medem a comunicação eficiente entre espalhador-ignorante ($\beta$) e espalhador-sufocador ($\alpha$), respectivamente. Cabe

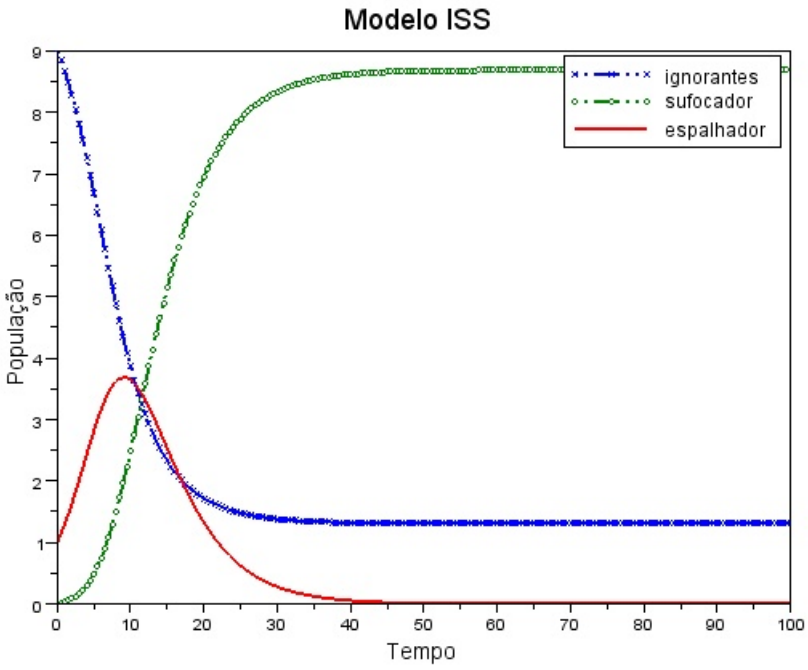


Figura 4.2: Variação temporal do número de ignorantes $I(t)$, do número de espalhadores $S(t)$ e do número de sufocadores $R(t)$.

ressaltar que os valores dessas taxas foram escolhidos arbitrariamente na literatura.

## 4.3 Modelo SIR com Dois Níveis de Atitude.

O modelo SIR (Suscetível, Infectado, Recuperado) com dois níveis de atitude foi apresentado por Pachi (PACHI, 2006), baseado no modelo compartimental SIR de Epidemiologia, que descreve a dinâmica da transmissão de ideias e investiga o espalhamento delas na população.

A população total $N$ é subdividida em classes de indivíduos "suscetíveis a informação" ($S$), "infectados" ($I$) e "recuperados" ($R$) a propagação de informação. Cada uma dessas classes será ainda subdividida em dois grupos: um grupo é formado pelos indivíduos mais ativos, enquanto que no outro grupo, os indivíduos passivos a informação. Assim, $N = S1 + I1 + R1 + S2 + I2 + R2$. Com $S_1$, $I_1$ e $R_1$ indivíduos ativos, mais receptíveis a informação e $S_2$, $I_2$ e $R_2$ os indivíduos passivos, mais resistentes à informação.

A dinâmica da transmissão da informação será então representada pelo seguinte diagrama:

Onde:

• $\beta_i$ $(i = 1, 2, 3, 4)$ - são os coeficientes de transmissão da informação entre os indivíduos "infectados" e "suscetíveis";

• $\gamma$ - é a taxa de recuperação ou de "esquecimento";

• $\lambda$ - é a taxa de transmissão da informação em massa ($\lambda^{-1}$ representa o

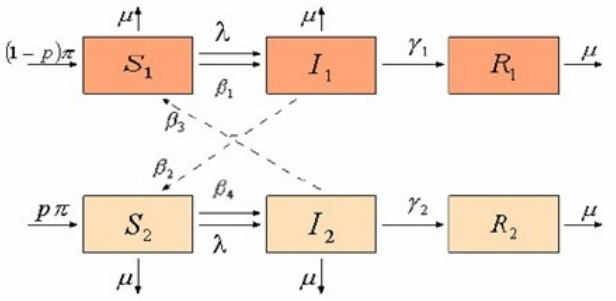


Figura 4.3: Diagrama de fluxo do Modelo de Transmissão de Informação. As setas pontilhadas indicam apenas o contato entre os indivíduos suscetíveis e infectados PACHI (2006).

período de duração da campanha educativa);

- $\mu$ - é a taxa de mortalidade natural;

- $p$ - é a proporção de indivíduos susceptíveis que entram no sistema $0 < p < 1$;

- $\pi$ - é o fluxo de entrada de indivíduos suscetíveis;

Considerou-se também duas possíveis formas da informação ser transmitida e então assimilada pelos indivíduos suscetíveis. Na primeira, os indivíduos suscetíveis podem ter acesso à nova ideia pelo contato direto com os indivíduos infectados, que já reconheceram e aceitaram a novidade. Na segunda situação, assume-se que os indivíduos suscetíveis são submetidos a uma campanha educativa, podem reconhecer e aceitar a nova ideia divulgada por qualquer meio de comunicação em massa e a transmissão da informação ocorre a uma taxa constante (PACHI, 2006).

Baseado nas suposições descritas anteriormente, a dinâmica do modelo que é dada pela Figura 4.3, pode ser então descrita pelo seguinte sistema de equações

diferencias ordinárias:

$$\frac{dS_1}{dt} = (1-p)\pi - \lambda S_1 - \beta_1 S_1 I_1 - \beta_3 S_1 I_2 - \mu S_1 \tag{4.10}$$

$$\frac{dS_2}{dt} = p\pi - \lambda S_2 - \beta_4 S_2 I_2 - \beta_2 S_2 I_1 - \mu S_2 \tag{4.11}$$

$$\frac{dI_1}{dt} = \lambda S_1 + \beta_1 S_1 I_1 + \beta_3 S_1 I_2 - (\gamma_1 + \mu) I_1 \tag{4.12}$$

$$\frac{dI_2}{dt} = \lambda S_2 - \beta_4 S_2 I_2 + \beta_2 S_2 I_1 - (\gamma_2 + \mu) I_2 \tag{4.13}$$

$$\frac{dR_1}{dt} = \gamma_1 I_1 - \mu R_1 \tag{4.14}$$

$$\frac{dR_2}{dt} = \gamma_2 I_2 - \mu R_2 \tag{4.15}$$

Com as condições iniciais não negativas e $N(0) > 0$ e $N = S_1 + I_1 + R_1 + S_2 + I_2 + R_2$.

Observe que $\frac{dN}{dt} = \pi - \mu N$, e que $\frac{dN}{dt} = 0$ se, e somente se, $\pi - \mu N = 0$, ou seja, quando $N = \frac{\pi}{\mu}$. Isto implica que $N(t) \longrightarrow \frac{\pi}{\mu}$ quando $t \longrightarrow \infty$. Portanto, se $N(t)$ for menor que $\frac{\pi}{\mu}$, a população aumenta, enquanto que para $N(t)$ maior que $\frac{\pi}{\mu}$, a população diminui até eventualmente atingir $\frac{\pi}{\mu}$, para todo $t$ suficientemente grande. Como a população total é constante, Pachi (PACHI, 2006) admitiu $\pi = \mu N$.

Pachi em (PACHI, 2006) propôs um modelo simplificado considerando uma situação alternativa, na qual a dinâmica da transmissão da novidade passa a ser afetada por uma consideração especial do comportamento dos indivíduos pertencentes à classe $I_2$. Pachi supôs que estes indivíduos, diferentemente dos indivíduos do grupo $I_1$, não possuem habilidades de bons comunicadores e não conseguem (ou não têm interesse) em divulgar as novidades.

Neste contexto, os indivíduos $I_2$ transmitem a inovação apenas aos indivíduos $S_1$, que por serem mais receptivos têm mais facilidade para aceitarem uma nova ideia e, portanto, teoricamente não se exige muita argumentação para persuadí-los. Em termos matemáticos, desconsideramos o contato entre os indivíduos $I_1$ e $S_2$, admitindo a taxa de contato $\beta_4 = 0$ no sistema de Equações (4.10 - 4.15).

Em contrapartida, as características das outras subpopulações e o comportamento mais e menos contumaz às novidades se mantêm em todas as outras classes. Assim, as duas possibilidades de transmissão da inovação ainda são: campanha educativa e contato direto com os indivíduos que reconhecem e aceitam a novidade.

A solução exata e a análise teórica deste modelo são encontradas na dissertação de Pachi em (PACHI, 2006).

A solução numérica com as condições iniciais e os parâmetros utilizados no sistema foram baseados no modelo apresentado por Pachi em PACHI (2006). Assim, consideramos as condições iniciais $S_1 = 0.30$, $S_2 = 0.25$, $I_1 = 0.20$, $I_2 = 0.5$, $R_1 = 0.10$ e $R_2 = 0.10$; os parâmetros $\beta_1 = 0.003$, $\beta_2 = 0.002$, $\beta_3 = 0.003$, $\beta_4 = 0.002$, $p = 0.2$, $\mu = 0.01388$, $\pi = 0.01388$, $\lambda = 0.8$, $\gamma_1 = 0.125$ e $\gamma_2 = 0.1666$. O gráfico obtido da simulação numérica está representado na Figura 4.4. Cabe enfantizar, que os valores dos parâmetros foram colocados de forma arbitrária no modelo proposto por Pachi em (PACHI, 2006).

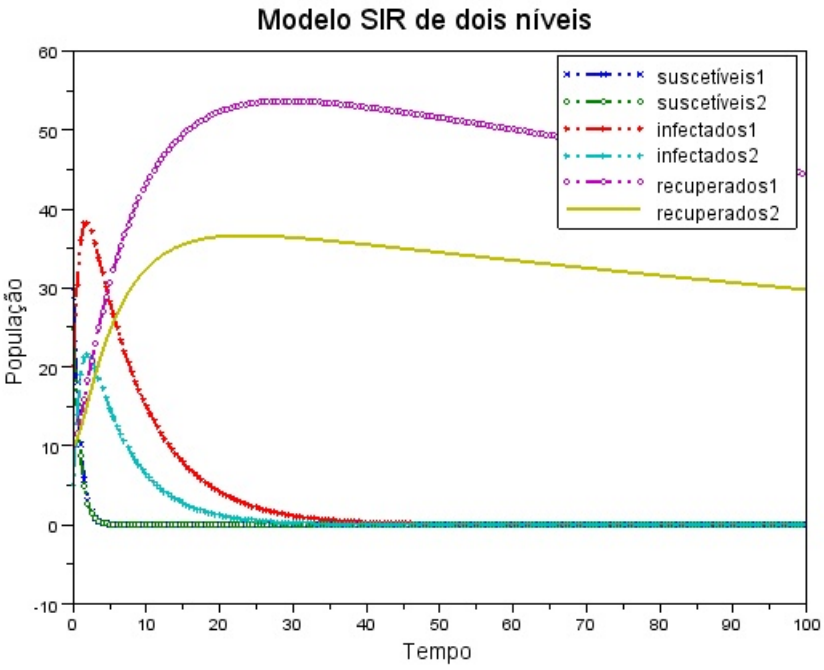


Figura 4.4: Dinâmica do sistema.

As análises da dinâmica do espalhamento de idéias numa população por meio da simulação numérica foram feita por Clarice, ela conseguiu interpretar as implicações desses resultados, associando-os a transmissão de informações por campanhas educativas e também por rumores. Clarice admite em sua dissertação que os rumores são elementos que intensificam a transmissão da informação e desta forma podem contribuir para melhorar a eficácia na propagação de idéias numa população.

## 4.4 Modelo SEIR para a propagação de informação.

O modelo SEIR (Suscetível - Exposto - Infectado - Recuperado) para propagação boca-a-boca da informação do Flickr (site da web de hospedagem e partilha de imagens fotográficas, e eventualmente de outros tipos de documentos gráficos, como desenhos e ilustrações, além de permitir novas maneiras de organizar as fotos e vídeos), foi apresentado por Cha et al., (CHA et al., 2012).

Cha et al. em (CHA et al., 2012), investigamos as possíveis razões para o atraso na propagação boca-a-boca das informações dos usuários do Flickr. Cha et al. em (CHA et al., 2012), identificaram dois fatores que pode alterar os espalhadores da informação numa rede social: a explossão de usuário no tempo de conexão e o envelhecimento do conteúdo. Estudaram o impacto desses fatores usando um modelo epidemiológico que foi adaptado para permitir a investigação da velocidade de propagação de boca-a-boca. A simulação mostrada por eles, diz que os dois fatores podem explicar os padrões observado nos dados reais e ajudar a compreender como esses fatores afetam a capacidade da rede para divulgar informações de forma rápida e amplamente.

Considere-se uma população de utilizadores de $N$ no tempo $t$, em que $S$, $E$, $I$ e $R$ são os usuários, suscetíveis, exposto, infectados e recuperado, respectivamente,

e $N = S + E + I + R$.

A dinâmica da transmissão da informação será então representada, pelo seguinte diagrama:

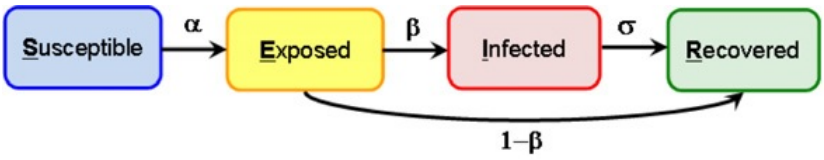


Figura 4.5: Fluxograma do modelo SEIR de propagação boca-a-boca.

onde:

• $\alpha$ - é a probabilidade de um utilizador suscetível, com a pelo menos um amigo infectado, ser exposto ao conteúdo.

• $\beta$ - é a probabilidade de que um usuário exposto favorito marcar a foto e torna-se infectado.

Eles admitiram que os usuários expostos que não marca a foto favorita, não marcará mais, assim, tornar-se imune à foto. Como resultado, os utilizadores expostos têm probabilidade $1 - \beta$ para passar para o estado recuperou.

• $\sigma$ - é a probabilidade de que um utilizador remova a foto infectada de sua lista de favoritos e, conseqüentemente, se recupera.

Baseado nas suposições descritas anteriormente, a dinâmica do modelo é dada pela Figura 4.5, pode ser então descrita pelo seguinte sistema de equações de diferenças:

$$S_{t+1} = S_t - \alpha S_t I_t \tag{4.16}$$

$$E_{t+1} = E_t + \alpha S_t I_t - \beta E_t - (1-\beta) E_t \tag{4.17}$$

$$I_{t+1} = I_t + \beta E_t - \sigma I_t \tag{4.18}$$

$$R_{t+1} = R_t + (1-\beta) E_t + \sigma I_t \tag{4.19}$$

O artigo (CHA et al., 2012) utilizou-se de equações de difenças ao invés de equações diferenciais, a mudança de tempo contínuo para discreto não produz nenhum efeito sobre a natureza básica da análise dinâmica, apesar de alterar a sua formulação. Quando o tempo é considerado uma variável discreta, ou seja, $t$ assuma apenas valores inteiros, o conceito de derivada perde sua aplicabilidade e o padrão de mudanças da variável dependente é descrito pelas chamadas diferenças. Neste sentido, o modelo dinâmico que procura determinar a trajetória no tempo, é descrito pelas equações de diferenças.

O modelo foi testado utilizando dados reais de redes sociais. Foram usadas redes de diferentes tamanhos (isto é, que varia de 5000 a 1 milhão de utilizadores). A simulação funciona da seguinte forma: Dada uma topologia de rede, começaram com um único usuário infectado. Os restantes utilizadores estão no estado suscetíveis. Em cada unidade de tempo de simulação, os usuários que fazem login têm uma chance de passar de um estado para outro de acordo com a dinâmica do modelo proposto.

Para ilustrar a dinâmica do modelo, mostra um exemplo da simulação de funcionamento com base em uma rede de 5000 usuários. A Figura 4.6 ilustra a dinâmica de espalhamento de uma fotografia, utilizando os parâmetros do modelo de $\alpha = 0{,}2$, $\beta = 0{,}2$, $\sigma = 0$, e nenhum atraso fatores. Neste cenário, pressupõe-se que todos os usuários entrem no Flickr, uma vez a cada dia. A simulação mostra que o número de utilizadores aumenta rapidamente e todos os usuários tornam-se infectados ou recuperados. Ao encontrar uma foto (ou seja, quando um amigo tem o favorito marcado em qualquer foto), os usuários têm uma chance de 20% da marcação favorito da foto. Consequentemente, o número de utilizadores infectados rapidamente aumenta e estabiliza a epidemia após 15 dias.

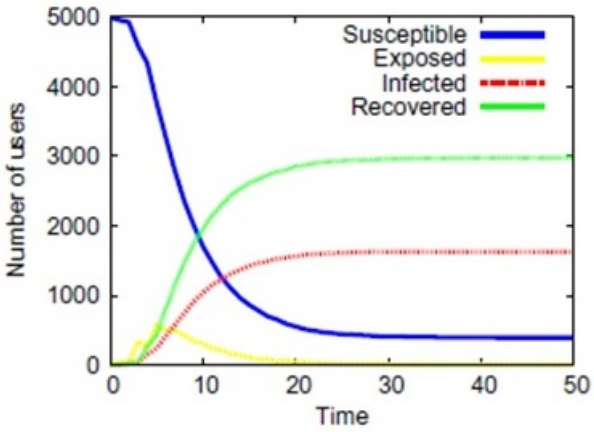


Figura 4.6: Dinâmica modelo SEIR de propagação boca-a-boca apresentado em CHA et al. (2012).

As conclusões feitas pelos autores foram que esta simulação de base, apresentou um padrão semelhante em uma variedade de tamanhos de rede e valores para o $\beta$ parâmetro. Para redes maiores, mais usuários permanecem no estado suscetível quando o padrão de epidemia estabiliza. Quando diminuir o valor de $\beta$, o número do total infectado utilizadores diminui.

No artigo não ficou claro como os autores retiraram as taxas apresentadas no modelo dos dados da rede social Flikr.

## 4.5 Considerações.

Ao estudarmos os modelos, para propagação da informação, baseados nos modelos epidemiológicos, percebemos que não há muita diferença entre eles. A maioria dos modelos apresentam população constante, muitos paramêtros de serem estimados e comparando o comportamento das dinâmicas dos modelos pelos gráficos, observamos que eles possuem dinâmicas parecidas.

Contudo, os estudos de tais modelos para a propagação de informação tem se mostrado crescente sobre os processos de contágio que podem ser manifestados por propagação de boatos, vírus de computador, difusão de inovações, revoltas políticas e disseminação de doutrinas religiosas. Podendo assim, ser uma ferramenta de importância significativa para: empresas, autoridades públicas, campanhas educativas; entre outros.

Percebendo o potêncial da modelagem matemática e a não consagração de nenhum modelo em especial dentro do universo da informação, como os apresentados neste capítulo, decidimos explorar no próximo capítulo o modelo mais simples, que tem a vantagem de trazer poucas taxas a serem estimadas e que apresenta mais de uma dinâmica. Além de permitir, que a população total deixe ser constante, o que, em muitas situações a população onde a informação pode se propagar, cresce, deixando o modelo mais próximo da realidade.

# 5 ADAPTAÇÃO DO MODELO SUSCETÍVEL - INFECTADO COM DINÂMICA VITAL PARA PROPAGAÇÃO DE INFORMAÇÃO

O modelo SI de epidemiologia foi escolhido para ser adaptado para um modelo de propagação de informação. A população total é dividida em suscetíveis ($S$) e infectados ($I$), usamos apenas dois grupos possíveis, porém permitindo que a população total varie. Dessa forma, o modelo é simples, porém apresenta mais de uma dinâmica. Os indivíduos que não ouviram a informação ou não a usam de forma ativa são, portanto, suscetíveis de serem informados e de se tornarem sujeitos ativos. Infectados são os indivíduos que possuem a informação e que de forma ativa estão espalhando a notícia.

O comportamento dinâmico do processo de transmissão de informação acontece quando uma pessoa suscetível encontra um infectado (Princípio da ação das massas), que é transformado em um novo infectado com probabilidade $\beta$ ($\beta > 0$) . Os infectados têm "morte" induzida pela informação, ou seja, devido a um processo de esquecimento porque a informação perdeu o seu valor de notícia com constante de proporcionalidade $\phi$ ($\phi > 0$).

Os indivíduos removidos de cada classe por "morte natural", por exemplo, o cancelamento da conta em uma das mídias sociais se o meio de propagação da informação for a mídia social, com taxa proporcional ao tamanho da classe com constante de proporcionalidade $\mu$ ($\mu > 0$). Existe também um fluxo de entrada de indivíduos suscetíveis na comunidade que representa os "nascimentos" com constante de proporcionalidade $b$ ($b > 0$). "Nascimentos" podem ocorrer em toda população (suscetíveis e infectados) ou apenas na população de suscetíveis. Analisaremos apenas a segunda

opção.

A dinâmica do modelo é representada pelo fluxograma (Figura 5.1) que corresponde ao seguinte sistema de equações:

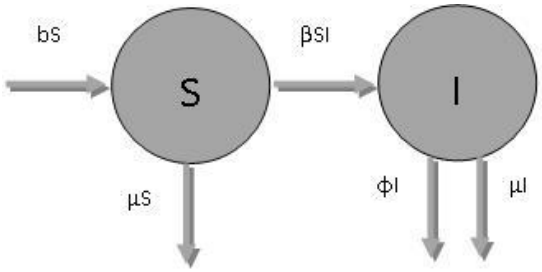


Figura 5.1: Fluxograma representando o modelo SI com fluxo de entrada diferente do fluxo de saída, onde apenas $S$ contribui para fluxo de entrada.

$$\frac{dS}{dt} = S(b - \beta I - \mu) \tag{5.1}$$

$$\frac{dI}{dt} = I(\beta S - \phi - \mu) \tag{5.2}$$

com condições iniciais $I(0) = I_0$ e $S(0) = S_0$.

Somando-se as Equações (5.12) e (5.13), obtemos:

$$\frac{dN}{dt} = bS - \phi I - \mu N, \tag{5.3}$$

o que nos mostra que neste modelo a população total não é necessariamente constante. Interessante perceber que esse modelo permite que a população total deixe de ser constante, assim permitindo que a população cresça.

Para ocorrer a epidemia, temos que ter um crescimento da quantidade de infectados por tempo, o que significa $\frac{dI}{dt} > 0$. Assim o lado esquerdo da Equação (5.2) torna-se $I(\beta S - \phi - \mu) > 0$, já que $I > 0$ então $\frac{\beta S}{\mu+\phi} > 1$. Esta fração representa o número médio de novas infecções causadas por um indivíduo infectado, durante seu período de infecciosidade. Assim, define-se um parâmetro, amplamente utilizado nos modelos matemáticos de epidemiologia, para avaliar estratégias de erradicação ou controle de uma doença numa população, o número reprodutivo básico, identificamos que o $(R_0)$ é o valor desta fração, no início da infecção, quando toda população é suscetível, isto é, quando $S = N(0) = N_0$. Portanto $R_0 = \frac{\beta N_0}{\mu+\phi}$.

Em geral, quando $R_0 < 1$, a doença desaparece. Quando $R_0 > 1$, haverá uma epidemia, ou seja, a doença infecciosa e transmissível ocorrerá numa comunidade ou região e pode se espalhar rapidamente entre as pessoas de outras regiões. Quando $R_0 = 1$, a doença irá tornar-se endêmica, ou seja, uma doença localizada em um espaço limitado, podendo se propagar ou não (CODECO CLAUDIA E COELHO, 2008).

Esta definição foi adaptada para $R_0^*$, e será definido como o número médio de indivíduos que adotam uma nova ideia, gerado por um único indivíduo infectado por ela e introduzindo numa população inteiramente suscetível à novidade. Para $R_0^* < 1$ tem-se a ausência da propagação da informação na população. Quando $R_0^* > 1$ a informação "invade" a população dando origem a sua propagação e para $R_0^* = 1$, pode ocorrer ou não a propagação da informação (LOPEZ et al., 2002).

Assim, de acordo com as definições acima, concluímos que o valor da fração no início da infecção é $R_0^*$ , quando toda população é suscetível, isto é, quando $S = N(0) = N_0$. Portanto $R_0^* = \frac{\beta N_0}{\mu+\phi}$.

## 5.1 Pontos de Equilíbrio e Análise de Estabilidade.

Faremos uma análise matemática que tem o objetivo de prever em tempos futuros como a população se comportará, ou seja, qual dinâmica ela atingirá. E ainda identificar casos do $R_0^*$ para verificarmos se a informação "invadirá"ou não a população.

O pontos de equilíbrio $(\widehat{S}, \widehat{I})$, obtidos a partir das Equações (5.1) e (5.2), tais que $\frac{dS}{dt} = 0$ e $\frac{dI}{dt} = 0$, são:

o equilíbrio com extinção da espécie, ou seja, onde não acontece mais a propagação da informação:

$$(\widehat{S_1},\ \widehat{I_1}) = (0, 0) \tag{5.4}$$

esse ponto de equilíbrio é atingido depois de algum tempo, já que a população não pode ser totalmente nula no tempo inicial.

E o equilíbrio endêmico, ou seja, onde há propagação da informação em um espaço limitado e de duração contínua:

$$(\widehat{S_2},\ \widehat{I_2}) = (\frac{\mu + \phi}{\beta}, \frac{b - \mu}{\beta}) \tag{5.5}$$

desde que $b - \mu > 0$, isto é, $b > \mu$. Se $b < \mu$, será impossível que a população

cresça, pois a Equação (5.3) será sempre decrescente e apenas o equilíbrio com extinção da espécie será possível.

A análise da estabilidade destes equilíbrios serão efetuadas através da linearização do sistema (BOYCE; DIPRIMA, 1995).

Linearizando o sistema de Equações (5.1) e (5.2), em torno do equilíbrio $(\widehat{S},\ \widehat{I})$, obtemos a seguinte matriz Jacobiana.

$$J(\widehat{S},\widehat{I}) = \begin{pmatrix} b-\beta\widehat{I}-\mu & -\beta\widehat{S} \\ \beta\widehat{I} & \beta\widehat{S}-\phi-\mu \end{pmatrix} \tag{5.6}$$

Da análise dos autovalores ($\lambda_1$ e $\lambda_2$) na matriz Jacobiana (5.6), em cada um dos pontos de equilíbrios (5.4) e (5.5), resulta a classificação de cada um, apresentada na Tabela 5.1:

Tabela 5.1: : Estabilidade de cada um dos pontos de equilíbrios (5.4) e (5.5).

| Parâmetro | Equilíbrio | Estabilidade |
|---|---|---|
| $b < \mu$ | $(\widehat{S_1},\ \widehat{I_1})$ | estável ($\lambda_1 < 0$ e $\lambda_2 < 0$) |
| | $(\widehat{S_2},\widehat{I_2})$ | instável ($\lambda_1 > 0$ e $\lambda_2 < 0$) |
| $b > \mu$ | $(\widehat{S_1},\ \widehat{I_1})$ | instável ($\lambda_1 > 0$ e $\lambda_2 < 0$) |
| | $(\widehat{S_2},\widehat{I_2})$ | indeterminado ($\lambda_1 = \mu i$ e $\lambda_2 = -\mu i$)(*) |
| $b = \mu$ | $(\widehat{S_1},\ \widehat{I_1})$ | caso degenerativo ($\lambda_1 = 0$ e $\lambda_2 < 0$)(**) |
| | $(\widehat{S_2},\widehat{I_2})$ | caso completamente degenerativo ($\lambda_1 = 0$ e $\lambda_2 = 0$)(***) |

(*) Inderteminado para o sistema não linear, mas nos leva a um sistema estável nas proximidades do ponto $(\widehat{S_2},\widehat{I_2})$ quando linearizamos o sistema;

(**) São pontos fixos associados as soluções constantes;

(***) Todas as soluções são constantes, com todos os pontos do plano sendo pontos fixos do sistema.

Concluímos, portanto, que se o fluxo de entrada de indivíduos suscetíveis for memor que o fluxo de saída das classes suscetível e infectado ($b < \mu$), o sistema tenderá ao único equilíbrio viável que é o da extinção da espécie. Por outro lado, se $b > \mu$, este equilíbrio é instável e nada podemos concluir a respeito do equilíbrio endêmico. E, para $b = \mu$, um caso degenerativo (ROSA, 2009), quando pelo menos um dos autovalores é nulo, assim, nada podemos concluir para a estabilidade dos pontos equilíbrios do ponto de vista linear.

Como a análise de estabilidade não é conclusiva quando pelo menos um autovalor tem parte real nula, o estudo a ser apresentado a seguir se restringirá ao caso em que ambos os autovalores tem parte real não nula. Calculando a solução geral do sistema linearizado, cuja matriz Jacobiana apresentada (5.6), em torno de cada um dos pontos pontos equilíbrios como feito em (ROSA, 2009), encontramos:

a) próximo ao equilíbrio $(\widehat{S}_1, \widehat{I}_1) = (0, 0)$:

$$S(t) = C_1 e^{(b-\mu)t} \; e \; I(t) = C_2 e^{-(\phi+\mu)t} \tag{5.7}$$

A partir da solução da Equação (5.7) observamos que depende apenas do sinal $b - \mu$, visto que $\mu + \phi$ é sempre positivo.

Se $b < \mu$ tem-se:

$$\lim_{t\to\infty}(S(t), I(t)) = (0, 0), \tag{5.8}$$

concluímos então, que se trata de um equilíbrio do tipo nó estável.

Se $b > \mu$, o limite (5.8) verifica-se apenas quando $C_1 = 0$, o que indica equilíbrio do tipo sela (instável).

As conclusões às quais chegamos acima, a partir da solução (5.7), com relação a estabilidade e do tipo de equilíbrio do ponto $(\widehat{S}, \widehat{I}) = (0, 0)$, estão de acordo com os resultados encontrados na Tabela 5.1 para este ponto.

b) próximo ao equilíbrio $(\widehat{S}_2, \widehat{I}_2) = (\frac{\mu+\phi}{\beta}, \frac{b-\mu}{\beta})$:

$$S(t) = \frac{\mu+\phi}{\beta} + C_3 s e^{st} - C_4 s e^{-st} \text{ e } I(t) = \frac{b-\mu}{\beta} + C_3(b-\mu)e^{st} + C_4(b-\mu)e^{-st}, \quad (5.9)$$

onde:

$$s = \sqrt{(\mu - b)(\phi + \mu)}$$

O estudo da estabilidade deste equilíbrio depende dos valores do argumento $s$ das funções exponencias envolvidas em (5.9), que por sua vez, dependem apenas do valor de $\mu - b$, uma vez que $\phi + \mu$ é sempre positivo.

Se $b < \mu$, tem-se $s$ real positivo, o

$$\lim_{t\to\infty}(S(t), I(t)) = (\frac{\mu+\phi}{\beta}, \frac{b-\mu}{\beta}), \quad (5.10)$$

verifica-se apenas quando $C_3 = 0$, o que indica um equilíbrio do tipo sela (instável).

Se $b > \mu$, tem-se $s$ imaginário puro, o que corresponde a um equilíbrio do tipo centro.

As conclusões acima, a partir da solução (5.9), ratificam os valores encontrados na Tabela 5.1 para a estabilidade e o tipo de equilíbrio do ponto $(\widehat{S}_2, \widehat{I}_2)$.

## 5.2 Resultados Qualitativos.

Utilizou-se o software Scilab 5.3.3 para os resultados numéricos com o objetivo de associá-lo ao estudo qualitativo do modelo analisado anteriormente e ver algumas dinâmicas para situações diferentes.

Os resultados encontrados simulam o comportamento da dinâmica de transmissão de informação por meio da propagação de ideias numa população suscetível às novidades.

Nesta seção, os parâmetros foram escolhidos de forma mostrar os diferentes comportamentos qualitativos apresentados pelas soluções do sistema.

Para obter as soluções numéricas do sistema de Equações (5.12) e (5.13), utilizou-se o método de Runge-Kutta de quarta ordem do ambiente Scilab.

A simulação funciona da seguinte forma: começa-se com um único indivíduo infectado, os 500 indivíduos restantes estão no estado suscetível. Em cada unidade de tempo da simulação, o indivíduo que entra em contato com outro têm uma chance de passar de um estado para outro de acordo com a dinâmica do modelo proposto.

Na Figura 5.2, dados os valores de $\beta = 0.003$ (taxa de contato devido o encontro com um indivíduo infectado por unidade de tempo), $b = 0.02$ ( taxa de nascimento), $\mu = 0.15$ (taxa de mortalidade) e $\phi = 0.02$ (taxa de mortalidade devida a falta de interesse no assunto), onde $b < \mu$ e $R_0^* \simeq 8,84$, portanto, ocorre a epidemia, ou seja , acontece a propagação da informação. Pode-se observar que com o passar do tempo, $S(t) \to 0$ e $I(t) \to 0$, sem comportamento oscilatório. Portanto, o ponto de equilíbrio com extinção da espécie $(0, 0)$ é estável.

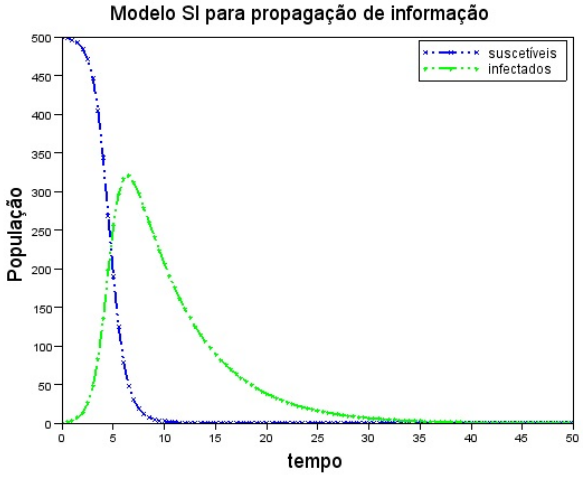


Figura 5.2: Variação temporal do número de suscetíveis $S(t)$ e do número de infectados $I(t)$, com $b < \mu$.

Este tipo de dinâmica sugere que a informação a ser propagada consegue atingir um número relevante de pessoas na rede social que foi disseminada, porém, com o passar do tempo perde seu valor de notícia. De forma qualitativa esta dinâmica pode ser vista no gráfico da Figura 5.3 que mostra a vida média de 1000 links populares para cada mídia social. Ele mostra quanto tempo um link está "vivo"antes das pessoas pararem de acessá-los ou propagá-los.

Na Figura 5.4, dados os valores de $\beta = 0.004$, $b = 1.32$, $\mu = 1.22$ $e$ $\phi = 1.15$, onde $b > \mu$ e $R_0^* \simeq 1.06$, portanto, ocorre a epidemia, ou seja, acontece a

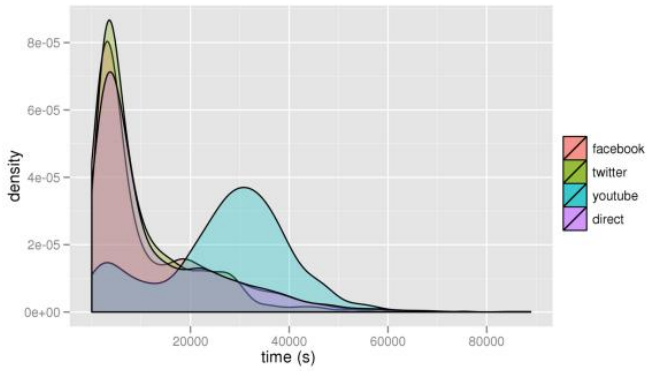


Figura 5.3: Distribuição da vida média do compartilhamento de links ao longo de quatro tipos diferentes de mídias sociais. (www.gestordemarkenting.com).

propagação de informação, observa-se que existe um comportamento periódico. A linearização em torno do ponto de equilíbrio $(\widehat{S_2}, \widehat{I_2})$ foi estável nas proximidades do ponto (com autovalores imaginários puros). Visualiza-se, porém, na Figura 5.3, que este ponto é um centro. Como apresenta movimentos oscilatórios, a previsão das variações cíclicas implica que uma queda no número de infectados não implica, necessariamente em término da infecção. Depois de certo tempo de diminuição no número de infectados, este começarão a aumentar, pois exite uma reposição de novos suscetíveis; e uma nova epidemia inicia.

Este tipo de dinâmica sugere que a informação a ser propagada precise constantemente de difusores para disseminá-las, por exemplo, campanhas como doação de sangue, de coleta de lixo nas praias, de vacinação, que como no gráfico da Figura 5.4, de forma qualitativa, pode ser comparado com a dinâmica descrita na Figura 5.5. Com o passar do tempo os indivíduos esquecem a informação e precisam de um incentivo para lembrá-los.

Na Figura 5.6, dados os valores $\beta = 0.0005,\ b = 0,\ \mu = 0\ e\ \phi = 0$, onde $b = \mu$. Com essas escolhas o modelo SI de epidemiologia sem dinâmica vital é recuperado como em (BUBNIAKOVA, 2007). É estável para o ponto de equilíbrio

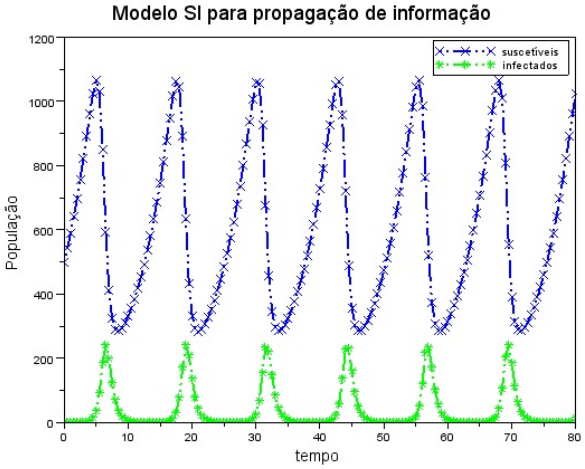


Figura 5.4: Variação temporal do número de suscetíveis $S(t)$ e do número de infectados $I(t)$, com $b > \mu$.

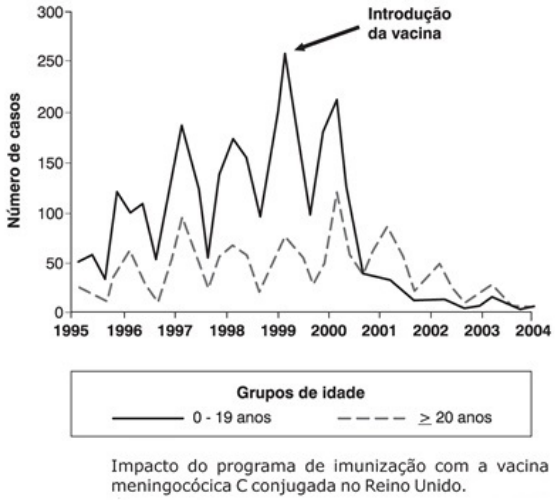


Figura 5.5: Variação temporal do número de infectados $I(t)$ após uma campanha de vacinação contra a meningite. (www.scielo.br).

$(0, N)$ e ocorre a propagação de informação.

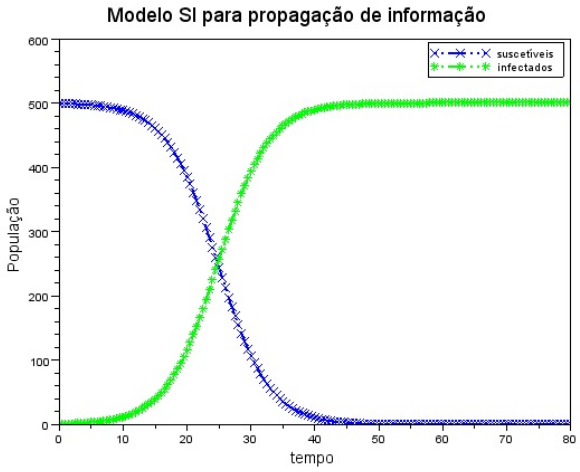


Figura 5.6: Variação temporal do número de suscetíveis $S(t)$ e do número de infectados $I(t)$, com $b = \mu = \phi = 0$.

Este tipo de dinâmica pode ser visto de forma qualitativa na Figura 5.7 a tendência por busca utilizando Facebook (Fonte: Googletrends), se considerarmos a curva como os indivíduos infectados.

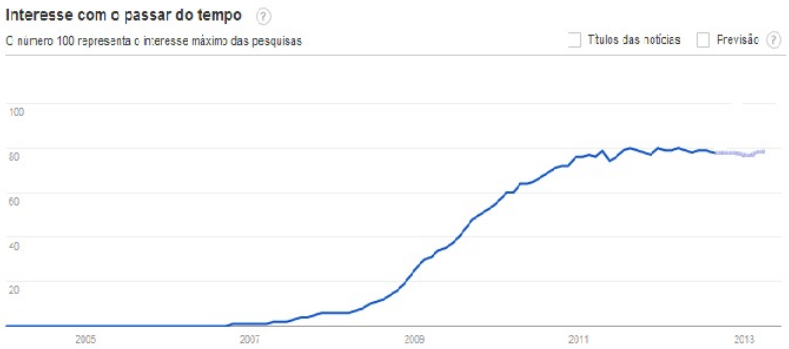


Figura 5.7: Variação temporal da tendência de busca por facebook. (www.googletrends.com acessado em 23 de julho de 2013 às 15 horas).

Nesta dissertação apresentamos algumas das possíveis dinâmicas, mas outras podem ser vistas quando variamos os valores dos parâmetros descritos no sistema

de Equações (5.1) e (5.2).

## 5.3 Metodologia para o Ajuste de Parâmetros.

Usamos o modelo mais simples utilizado em epidemiologia, pois ele contém apenas dois grupos de indivíduos com isso apresenta menos constantes, taxas que podem ser estimadas via dados retirados da informação que se deseja propagar. Lembrando que uma maneira seria estimar previamente as constantes que os modelos apresentam, mas percebemos a dificuldade de medir tais constantes, decidimos a partir do resultado estimar esses parâmetros.

No caso do SI, o ideal seria conhecer a taxa de contato devido o encontro com o indivíduo infectado por unidade de tempo ($\beta$) e usando-o no modelo. Porém, o que ocorre é que não sabemos o valor de $\beta$. Com os dados conseguimos identificar os indivíduos infectados ($I(t)$), uma alternativa é estimar $\beta$ a partir dos dados e por isso faremos um estudo de caso em uma das dinâmicas usando uma técnica do problema inverso.

O problema inverso é basicamente um problema de estimação de parâmetros, ou seja, dado uma solução de um sistema de EDO'S que depende de parâmetros desconhecidos queremos determinar os valores dos parâmetros. O problema inverso é uma alternativa para a calibração do modelo, a partir de uma série de dados que desejam descrevê-los, no qual os parâmetros têm de ser estimados a partir de amostras de dados. Matematicamente problemas inversos pertencem à classe de problemas mal-postos (CEZARO, 2008). No início deste século o matemático francês Jacques Hadamard definiu um problema bem-posto como sendo aquele cumpre três condições: existe solução, a solução é única e a solução tem uma dependência contínua (suave) com os dados de entrada. Assim, o problema é dito mal-posto se alguma

das condições acima não é satisfeita.

Existem alguns métodos de solução para problemas inversos (CAMPOS VELHO, 2010):

- Inversão direta;
- Decomposição em valores singulares;
- Mínimos quadrados e variantes (mínimos quadrados ponderados);
- Métodos de regularização;
- Métodos variacionais;
- Outros (molificação, métodos bayesianos, filtros digitais, redes neurais, etc).

Nesta dissertação encontra-se o método por mínimos quadrados.

### 5.3.1 Solução do Problema Inverso.

Um estudo inicial foi feito a respeito do problema inverso para a obenção da taxa necessária para o sistema de EDO's, material este que pode ser encontrado em (KUNZE; VISCAY, 1999), (CEZARO, 2008), (QUEIROZ LIMA, 2009).

Usamos o método de ajuste de mínimos quadrados para resolver o problema inverso. Utilizamos este método quando temos uma distribuição de pontos e queremos ajustar a melhor curva ao conjunto de dados.

Considerendo o problema de valor inicial:

$$\frac{dx}{dt} = f(x);\ x(t_0) = x_0$$

Em que $f(x)$ é dada como $f(x) = \sum_{k=1}^{m} \alpha_k \phi_k(x)$

A solução deste problema satisfaz a formulação da integral

$$x(t) = x_0 + \int_{t_0}^{t} f(x(s))ds = x_0 + \sum_{k=1}^{m} \alpha_k \int_0^t \phi_k(x(s))ds \tag{5.11}$$

Assim para qualquer solução de (5.11) temos que o desvio $\delta(x)$ deve ser *zero*, mas deve ser o menor possível quando consideramos uma aproximação $x(t)$.

$$\delta(x) = x(t) - x_0 - \sum_{k=1}^{m} \alpha_k \int_0^t \phi_k(x(s))ds$$

Escrevendo as funções de parâmetros que contempla os desvios, no método dos mínimos quadrados procuramos os parâmetros de forma que:

$$F(\alpha_1, ..., \alpha_m) = \int_0^T \parallel \delta(t) \parallel^2 dt \longrightarrow \text{mínimo.}$$

Trabalhamos com dados discretos, então: $[0, T] \rightarrow t_0, t_1, ..., t_n = n\Delta t; n = 0, ..., N$. Associando a discretização, temos os vetores $x_0, x_1 = x(t_1), ..., x_N = x(t_N)$.

Considerando, que o tempo inicial não variou, temos que calcular o desvio em cada tempo discreto; $t_0$ não tem desvio.

Assim:

$$\delta(x) = x_n - x_0 - \sum_{k=1}^{m} \alpha_k \int_0^{t_n} \phi_k(x(s))ds$$

Para calcular a integral $\int_0^{t_n} \phi_k(x(s))ds$ nos pontos discretos, vamos aproximar usando a regra dos trapézios.

$$\int_0^{t_n} f(\tau)d\tau \approx \sum_{j=0}^{n} w_j^n f(t_j) = \frac{\Delta t}{2}[f(t_0) + 2f(t_1) + ... + 2f(t_n - 1) + f(t_n)]$$

assim:

$$w_j^n = \Delta t \begin{cases} \frac{1}{2};\ \text{j} = 0 \text{ ou n.} \\ 1;\ \text{caso contrário.} \end{cases}$$

A expressão do desvio ficará assim:

$$\delta(x) \approx x_n - x_0 - \sum_{k=1}^{m} \alpha_k \sum_{j=0}^{n} w_j^n \phi_k(x_j)$$

Então a função:

$$F(\alpha_1, ..., \alpha_m) = \sum_{n=1}^{N} \parallel x_n - x_0 \sum_{k=1}^{m} \alpha_k \sum_{j=0}^{n} w_j^n \phi_k(x_j) \parallel^2$$

Ficamos com a expressão mais simples fazendo $\rho_k^n = \sum_{j=0}^{n} w_j^n \phi_k(x_j)$ e substituindo, obtemos:

$$F(\alpha_1, ..., \alpha_m) = \sum_{n=1}^{N} \parallel x_n - x_0 - \sum_{k=1}^{m} \alpha_k \rho_k^n \parallel^2$$

Derivando em relação para cada $\alpha_i$, temos que:

$$\frac{\partial}{\partial\alpha_i}\|x_n - x_0 - \sum_{k=1}^{m}\alpha_k\rho_k^n\|^2 = \frac{\partial}{\partial\alpha_i}(x_n - x_0 - \sum_{k=1}^{m}\alpha_k\rho_k^n, x_n - x_0 - \sum_{k=1}^{m}\alpha_k\rho_k^n) =$$

$$= -2(x_n - x_0 - \sum_{k=1}^{m}\alpha_k\rho_k^n, \rho_i^n).$$

Obtemos assim, as equações normais em que usamos o produto interno $(,)$.

$\sum_{n=1}^{N}(x_n - x_0 - \sum_{k=1}^{m}\alpha_k\rho_k^n, \rho_i^n) = 0$, i=1, ... , m.

Usando a notação de sistemas lineares:

$$\begin{cases} a_{11}\alpha_1 + a_{12}\alpha_2 + \ldots + a_{1m}\alpha_m = b_1 \\ a_{21}\alpha_1 + a_{22}\alpha_2 + \ldots + a_{2m}\alpha_m = b_2 \\ \qquad\qquad\qquad\vdots \\ a_{m1}\alpha_1 + a_{m2}\alpha_2 + \ldots + a_{mm}\alpha_m = b_m \end{cases}$$

onde:

$$a_{ij} = \sum_{j=1}^{m}[\sum_{n=1}^{N}(\rho_i^n, \rho_j^n)] \quad \text{e} \quad b_i = \sum_{n=1}^{N}(x_n - x_0, \rho_i^n).$$

Ao montar o sistema de equações, podemos observar que a matriz do sistema é simétrica e se as funções $\rho_{i,j}^n$ forem tais que os vetores $\rho_{i,j}^n$ resultantes forem linearmente independentes, o sistema admite uma única solução. A nossa expectativa é que este sistema tenha uma única solução. Então se o sistema tem uma solução única, esta solução é o ponto mínimo da função $F(\alpha_1, ..., \alpha_m)$.

Vamos analisar o nosso sistema de EDO'S, onde $\mu$, $\phi$ e $b = 0$:

$$\frac{dS}{dt} = S(b - \beta I - \mu) \tag{5.12}$$

$$\frac{dI}{dt} = I(\beta S - \phi - \mu) \tag{5.13}$$

o que retorna ao modelo SI mais simples.

Onde:

$$\phi_1 = SI \begin{pmatrix} 1 \\ -1 \end{pmatrix}; \alpha_1 = \beta$$

que será substituido em : $\frac{dx}{dt} = f(x) = \sum_{k=1}^{m} \alpha_k \phi_k(x)$.

## 5.4 Estudo de Caso.

Será apresentado um estudo de caso da aplicação da metodologia discutida, nos quais, dados da tendência de busca por Facebook obtidos no Google Trends são usados para se obter o parâmetro desejado. O Google Trends é uma ferramenta do Google que mostra os mais populares termos buscados em um passado recente. A ferramenta apresenta gráficos com a frequência em que um termo particular é procurado em várias regiões do mundo, e em vários idiomas. O eixo horizontal do gráfico representa tempo (a partir de algum tempo em 2004), e o vertical é com que frequência é procurado um termo, globalmente. Notícias relacionadas aos termos buscados são mostradas ao lado e relacionadas com o gráfico, apresentando possíveis

motivos para um aumento ou diminuição do volume de buscas. Os resultados obtidos não são fornecidos em termos de volume de pesquisa, mas sim dentro de um índice entre 0 a 100 em que 100 é o ponto no tempo em que a pesquisa teve o seu máximo atingido.

A simulação dos dados funciona da seguinte forma: atribuimos o índice ao número de pessoas onde começa-se com um único indivíduo infectado e os 99 indivíduos restantes estão no estado suscetível. Em cada unidade de tempo da simulação, o indivíduo que entra em contato com outro têm uma chance de passar de um estado para outro de acordo com a extração dos dados da Figura 5.7 que mostra somente a dinâmica dos infectados.

Usamos o software Scilab 5.3.3 para fazer o gráfico de dados, como ilustrado na Figura 5.8, cabe ressaltar que a curva dos indivíduos infectados que corresponde aos dados e a curva dos indivíduos suscetíveis é dado pela diferença entre a população total e os indivíduos infectados.

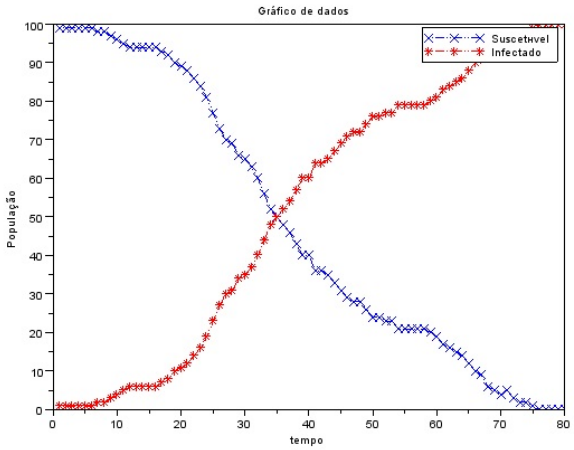


Figura 5.8: Variação temporal dos dados retirados do gráfico de tendência de busca por Facebook.

Podemos verificar que de forma qualitativa a Figura 5.8 representa a dinâmica do modelo SI mais simples de propagação de informação, Figura 5.6.

Aplicando o código de computador feito em Scilab 5.3.3 para resolução do problema inverso, ajustando a curva dos dados ao nosso modelo conseguimos prever qual será o valor de $\beta$ e o resultado encontrado foi $\beta = 4,8647 \text{ x } 10^4$.

Conforme o resultado que obtivemos, Figura 5.9, pode-se verificar que a dinâmica é correspondente ao gráfico dos dados. Para uma melhor comprensão das dinâmicas, os dois gráficos podem ser visualizados na Figura 5.10.

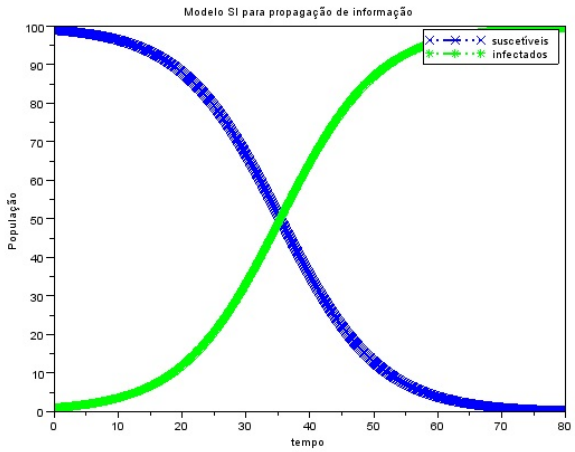


Figura 5.9: Variação temporal da propagação da informação com $\beta = 4,8647 \text{ x } 10^4$ estimado pelo problema inverso.

Com o modelo conseguimos não só prever a troca de interesse como manipular a informação de forma que essa troca possa acontecer em pouco tempo ou não. Como mostrado na Figura 5.11, onde aumentamos o valor de $\beta$ (taxa de contato devido ao encontro com um indivíduo infectado por unidade de tempo) e verificamos que a velocidade de propagação da informação dos dados calculados aconteceu em período menor. Já na Figura 5.13 diminuimos o valor de $\beta$ e consequentemente aumentou-se

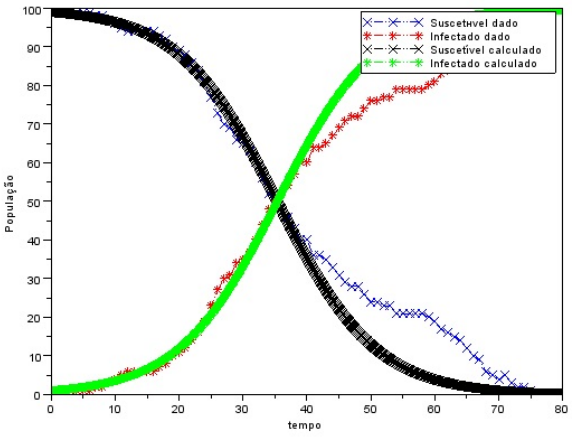


Figura 5.10: Gráfico de comparação - calculado pelo pelo problema inverso x dados retirados do Google Trends.

a velocidade de propagação da informação dos dados calculados.

Desta forma, o modelo de propagação de informação estará atento a capturar a capacidade de uma informação persuadir as pessoas.

Modelos deste tipo têm despertado especial interesse de grupos de estudos em Marketing empresarial que tem desenvolvido modelos com o objetivo de gerar predições a respeito da adoção de produtos ou para pesquisar sobre a tendência de opinião pública. Investimento em campanhas educacionais, principalmente as destinadas aos jovens, tem sido uma preocupação constante para as autoridades governamentais de todo o mundo. Apesar dos esforços constantes para melhorar a eficácia dessas campanhas, observa-se que em muitos casos, os resultados estão abaixo das expectativas.

A busca de um modelo que represente a dinâmica de uma propagação de informação ajuda na predição, mas é importante notar que complicar os modelos com

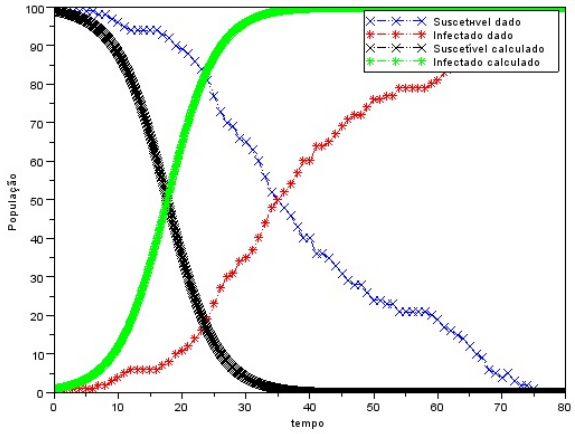


Figura 5.11: Gráfico de comparação com $(\beta = 4,8647 \text{ x } 10^4) * 2$.

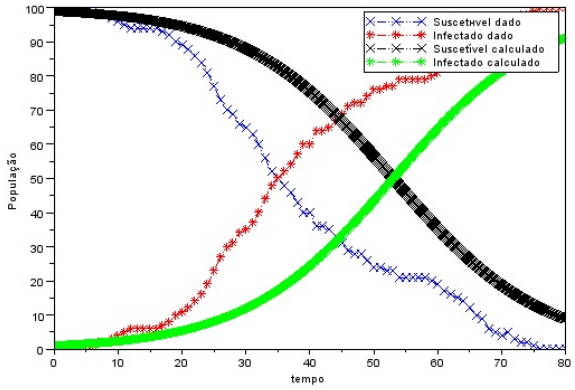


Figura 5.12: Gráfico de comparação com $(\beta = 4,8647 \text{ x } 10^4)/2$.

vista aproximá-los da realidade, por si só, não conduz as respostas adequadas aos problemas. Em geral, só cresce o número de perguntas. Ao tornar um modelo mais complexo, por um lado descreve-se melhor a realidade, por outro lado, aumenta-se o número de constantes do modelo, por isso, optamos em estudar o caso mais simples do modelo.

### 5.4.1 Outros Resultados.

Apresentaremos alguns resultados comparativos obtidos das simulções com o código da resolução do método do problema inverso para a obtenção da taxa de contato ($\beta$) e outras tendências de busca no Google Trends. As simulações seguem o mesmo que foi feito na seção 5.4, e, foram feitas visando testar e adquirir uma base de conhecimento sólida sobre o comportamento dinâmico do modelo.

Inicialmente, na Figura 5.13, são comparados os resultados da tendência de busca por Android no Google Trends. Os dados foram acessados em 15 de setembro de 2013, e, após a sua extração conseguimos fazer a estimação da taxa de contato ($\beta$). Verificamos que a dinâmica é correspondente ao gráfico de dados.

O gráfico ilustrado na Figura 5.14, mostra os resultados de comparação dos dados da tendência de busca por Google no Google Trends, que foram acessados em 15 de setembro de 2013 e o gráfico com valor da taxa de contato estimado pelo código. Verificamos que as dinâmicas são correspondentes.

Na Figura 5.14, comparamos os dados obtidos da tendência de busca por Yahoo com o gráfico obtido pelo código com o valor da taxa estimada. Os dados foram acessados em 15 de setembro de 2013 e podemos verificar que dinâmicas são correspondentes.

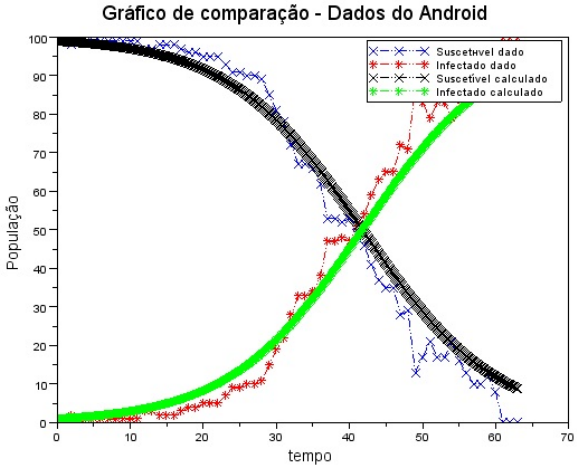


Figura 5.13: Gráfico de comparação dos dados calculados e retirados da tendência de busca por Android.

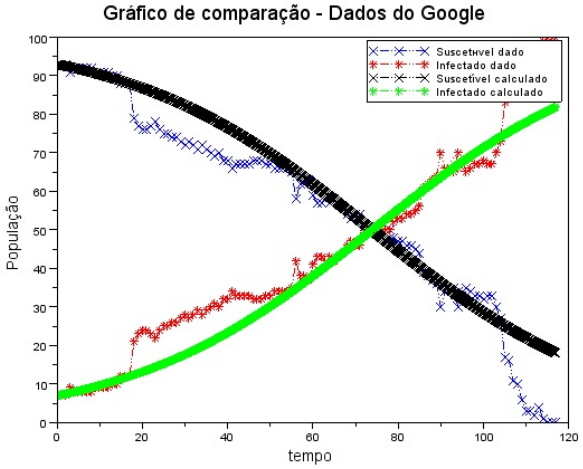


Figura 5.14: Gráfico de comparação dos dados calculados e retirados da tendência de busca por Google.

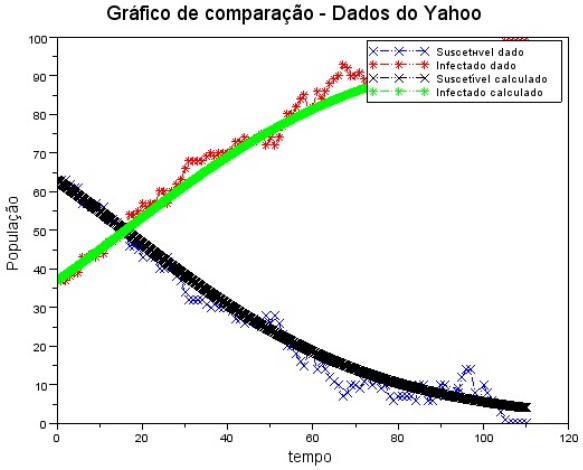


Figura 5.15: Gráfico de comparação dos dados calculados e retirados da tendência de busca por Yahoo.

Finalmente, para validarmos o código, retiramos os dados obtidos no modelo SI descrito na seção 3.1 e colocamos no código do problema inverso para estimar a taxa de contato. Obtemos os resultados da Figura 5.16. Assim verificamos que o código da resolução do problema inverso é válido.

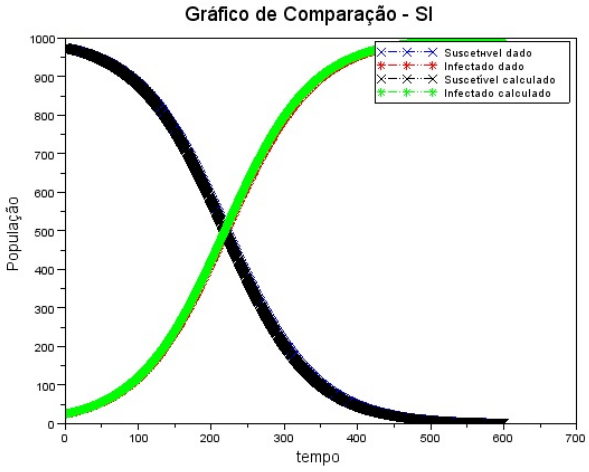


Figura 5.16: Gráfico de comparação dos dados calculados e retirados do gráfico da Figura 3.4.

# 6 CONCLUSÃO

Discutimos um modelo matemático simples que representa diferentes comportamentos na forma qualitativa da propagação da informação e que, por meio da analogia entre o espalhamento de idéias e doenças infecciosas, pode contribuir para compreensão das possíveis dinâmicas que envolvem a transmissão da notícia.

Interessante notar que a maior parte dos modelos de EDO's utilizados para a propagação de informação encontradas na literatura baseia-se em epidemiologia, contudo não são utilizados em situações reais, mesmo com os dados de como a informação se propaga em determinada população. A população desses modelos é de uma forma geral, considerada constante. Percebemos que essa hipótese deve ser alterada, dado que a maioria das populações onde se quer entender a propagação de informação varia ao longo do tempo. Percebemos ainda, que consideramos nascimentos e mortes na população, para que a população total deixasse de ser constante e a relação entre esses parâmetros proporciona diferentes dinâmicas no sistema.

O estudo da estabilidade dos pontos de equilíbrio nos permitiu perceber se o espalhamento da informação acontecerá ou não. E, através desta análise matemática podemos predizer em tempos futuros a dinâmica da população. Analisamos a dinâmica da disseminação da notícia numa população por meio de simulações numéricas e conseguimos interpretar as implicações desses resultados, associando-os a transmissão de informações em algumas situações diferentes.

Aplicamos o método dos quadrados mínimos na solução do problema inverso para EDO's, com objetivo determinar a taxa do modelo matemático mais simples a partir dos dados reais retirados da tendência de busca por Facebook, visto que,

extrair essas constantes dos informes são muito difíceis de medir na poulação mesmo quando se tem uma quantidade satisfatória de dados. Com o valor da taxa no modelo proposto, conseguimos não só estimar como a informação se desenvolve ao longo do tempo, e através das simulações podemos antecipar ou prolongar surtos do contágio social, e ajudar a desenvolver intervenções que controlem, ou espalhem, ou previnam, a propagação da informação.

Em tempos de redes sociais e maior contato, mesmo que virtual, entre as pessoas, entender e modelar o processo que descreve como a informação pode ser propagada é importante para diversos setores. Por exemplo, do ponto de vista comercial, entender a dinâmica de uma propaganda através de mais de um veículo de informação pode melhorar as vendas e diminuir investimentos. Outra situação, que de fato mais nos interessa, é usar esse conhecimento em campanhas educativas, com o objetivo de compreender o fenômeno e tornar uma campanha mais eficiente, ou seja, atingir muita gente com pouco gasto e provocar mudanças de comportamentos e atitudes.

O modelo apresentado não pretende de forma alguma cobrir todos os possíveis tipos de propagação de informação. Também não pretende espelhar todos os fatores relevantes para a dinâmica de qualquer tipo de informação. É apenas uma possível forma de enquadramento dos vários tipos de informação, uma espécie de primeira escolha, a partir da qual se poderão efetuar elaborações adicionais de modo a tornar o modelo mais próximo da realidade qualitativamente. Cobrir adequadamente realismo e viabilidade é o passo mais delicado em um processo de modelagem, por isso, o modelo matemático de propagação de informação apresentado foi mais simples e geral, que contemplou mais de uma dinâmica, que dependendo dos parâmetros utilizados pode ser usado em diferentes tipos de propagação de informação.

Acreditamos que este tipo de estudo tem importância significativa para a

pesquisa científica e estamos investigando a possibilidade da utilização desse estudo em outros dados conhecidos, como no contexto de aprendizagem de máquina, nos dados e entrega de conteúdo em redes sociais (EITAN ALTMAN FRANCESCO DE PELLEGRINI; JIMENEZ, 2013), (ALTMAN, 2014) e (ALTMAN, 2004), ao qual pretendemos analisar a eficiência de sua resposta em termos de coerência com os resultados conhecidos historicamente.

# REFERÊNCIAS

ALTMAN, E. An epidemic game between contents in a wireless networ. **Proc of IEEE IWCSC**, [S.l.], August 2004. hal-00690347, version 1.

ALTMAN, E. A stochastic game approach for competition over popularity in social networks. **Elsevier, special Issue on stochastic games**, [S.l.], v.1, n.1, 2014.

ALVARENGA, L. d. R. **Modelagem de epidemias através de modelos baseados em indivíduos**. 2008. Dissertação (Mestrado em Ciência da Computação) — Universidade Federal de Minas Gerais.

AMAKU, M. **Estudo de parâmetros epidemiológicos através de modelamento matemático**: aspectos estacionários, espaciais e temporais. 2001. Tese (Doutorado em Ciência da Computação) — USP- Instituto de Física - Departamento de Física Matemática.

BERNOULLI, D. **Essai d'une nouvelle analyse de la mortalité causée par la petite vérole et des advantages de l'inoculation pour la prévenir.** [S.l.]: Mémoires de Mathématiques et de Physique, 1760.

BILLINGS, L.; SPEARS, W. M.; SCHWARTZ, I. B. A unified prediction of computer virus spread in connected networks. **Physics Letters A**, [S.l.], 2002.

BOYCE, W. E.; DIPRIMA, R. C. **Equações Diferenciais Elementares e Problemas de Contorno**. [S.l.]: Editora Guanabara Koogan S. A., 1995.

BRAUER, F.; CASTILLO-CHAVEZ, C. **Mathematical Models in Population Biology and Epidemiology**. [S.l.]: Springer New York, 2012.

BUBNIAKOVA, L. **The Mathematics of infectious diseases**. 2007. Dissertação (Mestrado em Ciência da Computação) — Comenius University.

CAMPOS VELHO, H. F. de. **Problemas Inversos**: conceitos básicos e aplicações. [S.l.]: Instituto Nacional de Pesquisas Espaciais, Lab. de Computação e Matemática Aplicada, 2010.

RIO GRANDE, E. e Física da Universidade do (Ed.). **Problemas Inversos**: uma introdução. [S.l.]: FURG, 2008.

CHA, M. et al. Delayed information cascades in Flickr: measurement, analysis, and modeling. **Computer Networks**, [S.l.], 2012.

CODECO CLAUDIA E COELHO, F. **Introdução aos modelos de transmissão de doenças infecciosas**. [S.l.]: FIOCRUZ, 2008.

DIETZ, K.; HEESTERBEEK, J. A. P. Daniel Bernoulli's epidemiological model revisited. **Mathematical Biosciences**, [S.l.], v.180, 2002.

DODDS, P. S.; WATTS, D. J. Universal Behavior in a Generalized Model of Contagion. **Phys. Rev. Lett.**, [S.l.], 2004.

E., M.; F., R. A.; F., L. L. Modeling the spread of memes: how units of knowledge are transferred from brain to brain. **Journal of Information Science**, [S.l.], 2006.

EITAN ALTMAN FRANCESCO DE PELLEGRINI, R. E.-A. D. M.; JIMENEZ, T. Emergence of Equilibria from Individual Strategies in Online Content Diffusion. **NetSciCom**, [S.l.], 2013. Torino.

HETHCOTE, H. The Mathematics of Infectious Diseases. **SIAM Review**, [S.l.], v.42, n.4, p.599–653, 2000.

KERMACK, W. O.; MCKENDRICK, A. G. A Contribution to the Mathematical Theory of Epidemics. **Proceedings of the Royal Society of London. Series A**, [S.l.], v.115, n.772, p.700–721, 1927.

KUNZE, H. E.; VISCAY, E. R. **Solving Inverse Problems for ODEs using the Picard Contraction Mapping**. [S.l.]: Department of Aplied Mathematics - University of Waterloo, 1999.

LOPEZ, L. F. et al. Threshold conditions for infection persistence in complex host-vectors interactions. **Comptes Rendus Biologies**, [S.l.], 2002.

MARTIN NOWAK, R. M. M. **Virus Dynamics**: mathematical principles of immunology and virology. [S.l.]: Oxford University Press, 2000.

MASSAD, E. **Modelagem In**: forattini op. epidemiologia geral. 1th.ed. [S.l.]: Artes Médicas, 1996.

MIMS, C. E. A. **Microbiologia Médica**. [S.l.]: Monele Ltda, 2000.

MORENO, Y.; NEKOVEE, M.; PACHECO, A. F. Dynamics of rumor spreading in complex networks. **Phys. Rev. E**, [S.l.], 2008.

NAKAHARA, H. H. Avanços das redes sociais no Brasil. **IAB Brasil**, [S.l.], 2010.

NOWAK, M. **Evolutionary Dynamics**: exploring the equations of life. [S.l.]: Harvard University Press, 2006.

PACHI, C. G. d. F. **Modelo matemático para o estudo da propagação de informações por campanhas educativas e rumores**. 2006. Dissertação (Mestrado em Ciência da Computação) — Faculdade de Medicina da Universidade de São Paulo.

PEREIRA, M. G. **Epidemiologia - Teoria e prática**. [S.l.]: Guanabara Koogan S. A., 2000.

PIQUEIRA, J. R. C. Rumor Propagation Model: an equilibrium study. **Mathematical Problems in Engineering**, [S.l.], 2010.

PITTEL, B. On a Daley-Kendall Model of Random Rumours. **Journal of Applied Probability**, [S.l.], 1990.

QUEIROZ LIMA, R. de. **Identificação de Parâmetros pelos Métodos dos Mínimos Quadrados Não Linear**. [S.l.]: Pontífica Universidade Católica do Rio de Janeiro - PUC, 2009.

RAPOPORT, A. Spread of information through a population with socio-structural bias: i. assumption of transitivity. **The bulletin of mathematical biophysics**, [S.l.], 1953.

RIZZI, R. L. Resolução de Problemas e Modelagem Matemática. In: . [S.l.: s.n.]. p.66p.

ROSA, R. M. S. **Equações Diferenciais**. [S.l.]: IM, UFRJ, 2009.

SZABO, G.; A. HUBERMAN, B. o. Predicting the Popularity of Online Content. In: **Communications of the acm**. [S.l.: s.n.], 2010. v.53, n.8.

TSUJIGUCHI, A. C. **Um modelo de EDO aplicado à Influenza**. 2010. Dissertação (Mestrado em Ciência da Computação) — Universidade Estadual de Mato Grosso do Sul.

ZILL D.G.; CULLEN, M. **Equações Diferenciais**. 3.ed. [S.l.]: Makron Books, 2001. 434p. v.2.

# APÊNDICE A ANÁLISE DE ESTABILIDADE DOS EQUILÍBRIOS DE UMA EQUAÇÃO DIFERENCIAL ORDINÁRIA

Considerando a equação diferencial autônoma não linear

$$\frac{dx}{dt} = f(x) \tag{A.1}$$

e sendo, $\widehat{x}$ o valor de $x$, tal que $f(\widehat{x}) = 0$, isto é:

$$\left[\frac{dx}{dt}\right]_{x=\widehat{x}} = 0 \tag{A.2}$$

Dizemos que $\widehat{x}$ é um equilíbrio.

## A.1 Linha de fase.

Para construir e analisar a linha de fase de uma equação autônoma como em A.1 deve-se:

• resolver $f(x) = 0$ para determinar todas as soluções de equilíbrio e marcá-las numa linha de fase (eixo $x$ orientado);

• determinar o sinal de $\frac{dx}{dt}$ nos intervalos entre as soluções de equilíbrio sucessivas, e marcar setas horizontais apontando para direita quando $f(x) > 0$ e setas horizontais apontando para esquerda quando $f(x) < 0$;

• cada solução de equilíbrio pode ser classificado em: estável, instável ou semiestável, de acordo com a figura A.1:

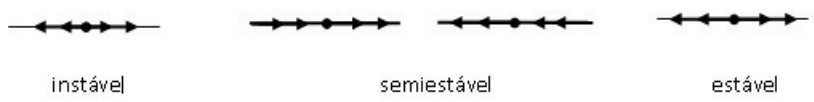


Figura A.1: Análise dos tipos de equilíbrio em uma linha de fase.

# APÊNDICE B ANÁLISE DE ESTABILIDADE DOS EQUILÍBRIOS DE UM SISTEMA BIDIMENSIONAL

## B.1 Linearização.

Linearização é a técnica que relaciona as soluções no caso não linear com as correspondentes soluções da aproximação linear do sistema em torno do equilíbrio em questão, para obter informações qualitativas a respeito do comportamento de longo prazo das soluções, a partir de um momento em que a solução esteja suficiente próxima do equilíbrio considerado.

O teorema de Linearização de Lyapunov-Poincaré nos auxilia no estudo de estabilidade dos pontos de equilíbrio.

Teorema 1: (Teorema da Linearização Lyapunov-Poincaré): Seja $f(x)$ um sistema de equações diferenciais de primeira ordem não-lineares continuamente diferenciável em uma vizinhaça de um ponto de equilíbrio $O$ e considere o sistema linearizado em torno de $O$ obtido pela expansão em série Taylor, isto é, $f(x) = Ax+B(x)$.

a) Então, se todos os autovalores da matriz quadrada $A$ possuírem parte real negativa, o ponto de equilíbrio $O$ será assintoticamente estável para o sistema não-linear.

b) Se pelo menos um autovalor $\Lambda$ de $A$ for tal que $Re(\Lambda) > 0$ (parte real positiva), o ponto de equilíbrio $O$ é instável para o sistema não-linear.

Dado o sistema:

$$\frac{dx}{dt} = P(x, y) \tag{B.1}$$

$$\frac{dy}{dt} = Q(x, y) \tag{B.2}$$

Que pode ser escrito em forma matricial:

$$\begin{bmatrix} \frac{dx}{dt} \\ \frac{dy}{dt} \end{bmatrix} \approx J(\widehat{x}, \widehat{y}) \begin{bmatrix} u \\ v \end{bmatrix} \tag{B.3}$$

onde definimos:

$u(t) = x(t) - \widehat{x}$ e $v(t) = v(t) - \widehat{y}$

e $J(\widehat{x}, \widehat{y})$ a matriz Jacobiana calculada no ponto $(\widehat{x}, \widehat{y})$:

$$J(\widehat{x}, \widehat{y}) = \begin{bmatrix} P_x(\widehat{x}, \widehat{y}) & P_y(\widehat{x}, \widehat{y}) \\ Q_x(\widehat{x}, \widehat{y}) & Q_y(\widehat{x}, \widehat{y}) \end{bmatrix} \tag{B.4}$$

Para verificar a estabilidade dos pontos de equilíbrio, deve-se analisar os autovalores $\Lambda_1$ e $\Lambda_2$ da matriz Jacobiana.

Os autovalores ($\Lambda$) da matrix Jacobiana $J(\widehat{x}, \widehat{y})$, são as soluções da equação algébrica:

$$\Lambda^2 - \Lambda T + D = 0 \tag{B.5}$$

onde $T$ e $D$ são o traço e o determinante da matriz $J(\widehat{x}, \widehat{y})$ , respectivamente.

Resolvendo a Equação (B.5) obtemos:

$$\Lambda_{1,2} = \frac{T \pm \sqrt{T^2 - 4D}}{2} \tag{B.6}$$

As raízes de (B.6) dependem do valor de $\Delta = T^2 - 4D$:

- se $\Delta > 0$ temos duas raízes reais e distintas, isto é, $\Lambda_1 ? \Lambda_2$;

- se $\Delta = 0$ temos duas raízes reais e iguais, isto é,$\Lambda_1 = \Lambda_2$;

- se $\Delta < 0$ temos raízes complexas com parte imaginária não nula.

Para classificar os pontos de equilíbrio podemos usar os critérios usados nos sistemas lineares, que estão apresentados na tabela A.2.

Se $D = 0$, temos uma das raízes da Equação (B.5) igual a zero. A outra raiz será positiva se $T > 0$ e negativa se $T < 0$. Esses casos são considerados raros e significam que existe uma região onde os pontos de equilíbrio não são isolados, e estão todos situados sobre uma reta. Um ponto de equilíbrio é isolado, se podemos determinar um círculo com centro dentro do qual não existem outros pontos de equilíbrio.

Tabela B.1: Propriedades de estabilidade dos pontos de equilíbrio de sistemas lineares.

| $\Delta = T^2 - 4D$ | Determinante | Traço | Tipo de ponto crítico | Estabilidade |
|---|---|---|---|---|
| $\Delta > 0$ | $D > 0$ | $T > 0$ | nó | instável |
| | $D > 0$ | $T < 0$ | nó | estável |
| | $D < 0$ | $T > 0$ ou $T < 0$ | ponto de sela | instável |
| $\Delta = 0$ | $D > 0$ | $T > 0$ | nó | instável |
| | $D > 0$ | $T < 0$ | nó | estável |
| $\Delta < 0$ | $D > 0$ | $T > 0$ | ponto espiral | instável |
| | $D > 0$ | $T < 0$ | ponto espiral | estável |
| | $D > 0$ | $T = 0$ | centro | estável |

# APÊNDICE C MÉTODOS DE RUNGE-KUTTA

Provavelmente um dos processos numéricos mais populares e, ao mesmo tempo, mais precisos para obter soluções aproximadas de equações diferenciais ordinárias é o método de Runge-Kutta de quarta ordem (ZILL D.G.; CULLEN, 2001).

Os métodos de Runge-Kutta (RK) possuem um erro local de truncamento de alta ordem. Essa propriedade tem um impacto direto no custo computacional do problema, pois os valores de $h$ estão diretamente relacionados ao número de iterações necessárias para a convergência da solução. Quanto menor for o valor do passo $h$, maior será o número de iterações necessárias para solucionar a equação para o mesmo intervalo $[t_0; t_f]$.

## C.1 Método de Runge-Kutta de Segunda Ordem (RK2).

$w_0 = y_0$

$k_1 = hf(t_i; w_i)$

$k_2 = hf(t_i + \frac{h}{2}; wi + \frac{1}{2}k_1)$

$w_{i+1} = w_i + k_2 + O(h^2)$

para cada $i = 1, 2, ..., N - 1$.

## C.2 Método de Runge-Kutta de Quarta Ordem (RK4).

$$w_0 = y_0$$

$$k_1 = hf(t_i; w_i)$$

$$k_2 = hf(t_i + \tfrac{h}{2}; wi + \tfrac{1}{2}k_1)$$

$$k_3 = hf(t_i + \tfrac{h}{2}; wi + \tfrac{1}{2}k_2)$$

$$k_4 = hf(t_i + h; wi + k_3)$$

$$w_{i+1} = w_i + \tfrac{1}{6}(k_1 + 2K_2 + 2k3 + k_4) + O(h^4)$$

para cada $i = 1, 2, ..., N - 1$. Esse método tem o erro local de truncamento $O(h^5)$ e o erro global que é formado pela acumulação dos erros de truncamento local a cada passo, juntamente com os erros de arredondamento, ou seja, é erro máximo do método no intervalo total da aproximação é $O(h^4)$ (**?**).

### C.2.1 Algoritmo - Método de Runge-Kutta de Quarta Ordem para os Sistemas de Equações Diferenciais.

O algoritmo abaixo implementa o método de Runge-Kutta de quarta ordem para os sistemas de valor inicial.

Para aproximar a solução do sistema de m-ésima ordem dos problemas de valor inicial de primeira ordem $u'_j = f_j(t; u_1; u_2; ...; u_m), t_0 \leq t \leq t_f$, com $u_j(t_0)$, com $u'_j(t_0) = \alpha_j$, para $j = 1, 2, ..., m$ em $(N + 1)$ números uniformemente espaçados no

intervalo $[t_0; t_f]$;

ENTRADA: pontos extremos $t_0$, $t_f$; número de equações $m$; inteiro $N$; condições iniciais $\alpha_1, ..., \alpha_m$.

SAÍDA: aproximações $w_j$ para $u_j(t)$ nos $(N+1)$ valores de $t$.

*Passo 1* - Faça

$h = \frac{(t_f - t_0)}{N}$;

*Passo 2* - para $j = 1, 2, ..., m$ faça $w_j = \alpha_j$;

*Passo 3* SAÍDA $(t; w_1; w_2; ...; w_m)$.

*Passo 4* - Para $i = 0, 1, 2, ..., N$ siga os *Passos 5-7.*

*Passo 5* - Para $j = 1, 2, ..., m$ faça

$k_{1,j} = hf_j(t; w_1; w_2; ...; w_m)$;

$k_{2,j} = hf_j(t + \frac{h}{2}; w_1 + \frac{k_{1,1}}{2}; w_2 + \frac{k_{1,2}}{2}; ....; w_m + \frac{k_{1,m}}{2})$;

$k_{3,j} = hf_j(t + \frac{h}{2}; w_1 + \frac{k_{2,1}}{2}; w_2 + \frac{k_{2,2}}{2}; ....; w_m + \frac{k_{2,m}}{2})$;

$k_{4,j} = hf_j(t + h; w_1 + k_{3,1}; w_2 + k_{3,2}; ...; w_m + k_{3,m})$;

$w_j = w_j + \frac{(k_{1,j} + 2k_{2,j} + 2k_{3,j} + k_{4,j})}{6}$;

*Passo 6* - Faça $t = t_0 + ih$;

*Passo 7* - SAÍDA $(t; w_1; w_2; ...; w_m)$.

*Passo 8* - PARE.