



UNIVERSIDADE FEDERAL DE SÃO CARLOS
CENTRO DE CIÊNCIAS EXATAS E TECNOLOGIA
DEPARTAMENTO DE MATEMÁTICA



JOÃO MARCOS CAVALCANTE PEDROSO

DISSEMINAÇÃO DA GONORREIA E COVID-19:
ESTUDO DE DOIS CENÁRIOS EPIDEMIOLÓGICOS

SÃO CARLOS – SP
2025

JOÃO MARCOS CAVALCANTE PEDROSO

DISSEMINAÇÃO DA GONORREIA E COVID-19:
ESTUDO DE DOIS CENÁRIOS EPIDEMIOLÓGICOS

Monografia apresentada como requisito parcial
para a obtenção do grau de Licenciatura em
Matemática na UFSCar.

Orientadora: Profa. Dra. Selma Helena de
Jesus Nicola

SÃO CARLOS – SP
2025



FUNDAÇÃO UNIVERSIDADE FEDERAL DE SÃO CARLOS

COORDENAÇÃO DOS CURSOS DE GRADUAÇÃO EM MATEMÁTICA - CCM/CCET

Rod. Washington Luís km 235 - SP-310, s/n - Bairro Monjolinho, São Carlos/SP, CEP 13565-905

Telefone: (16) 33518221 - <http://www.ufscar.br>

DP-TCC-FA nº 11/2025/CCM/CCET

Graduação: Defesa Pública de Trabalho de Conclusão de Curso

Folha Aprovação (GDP-TCC-FA)

FOLHA DE APROVAÇÃO

JOÃO MARCOS CAVALCANTE PEDROSO

DISSEMINAÇÃO DA GONORREIA E COVID-19: ESTUDO DE DOIS CENÁRIOS EPIDEMIOLÓGICOS

Trabalho de Conclusão de Curso

Universidade Federal de São Carlos – Campus São Carlos

São Carlos, 20 de fevereiro de 2025

ASSINATURAS E CIÊNCIAS

Cargo/Função	Nome Completo
Orientador	Selma Helena de Jesus Nicola
Membro da Banca 1	Vera Lúcia Carbone
Membro da Banca 2	Thaís Maria Dalbelo



Documento assinado eletronicamente por **Vera Lucia Carbone, Professor(a) do Ensino Superior**, em 16/04/2025, às 14:05, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do [Decreto nº 8.539, de 8 de outubro de 2015](#).



Documento assinado eletronicamente por **Selma Helena de Jesus Nicola, Professor(a) do Ensino Superior**, em 16/04/2025, às 17:02, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do [Decreto nº 8.539, de 8 de outubro de 2015](#).



Documento assinado eletronicamente por **Thais Maria Dalbelo, Professor(a) do Ensino Superior**, em 16/04/2025, às 19:55, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do [Decreto nº 8.539, de 8 de outubro de 2015](#).

Dedico este trabalho a todos que, de alguma forma, colaboraram para o meu desenvolvimento acadêmico, especialmente minha família, amigos, e minha orientadora.

AGRADECIMENTOS

Agradeço minha mãe e minha avó, por seu imenso apoio diante de todas as dificuldades que enfrentei. Agradeço também minha orientadora, que aprimorou de maneira incalculável este trabalho com suas sugestões, correções e incentivos.

RESUMO

Este trabalho apresenta modelos matemáticos para duas doenças que não conferem imunidade permanente após a infecção: a gonorreia e a COVID-19. No caso da gonorreia, são introduzidas equações diferenciais não lineares para descrever a dinâmica entre suscetíveis e infectadas numa determinada população. Analisamos estabilidade de pontos de equilíbrio e condições que levam à extinção ou persistência endêmica.

Posteriormente, abordamos a COVID-19 utilizando dados analisados de São Carlos, SP. Em vez de empregar um modelo compartimental clássico, recorreremos a um ajuste de curva logística para descrever o número acumulado de casos, discutindo como esse tipo de ajuste pode representar a evolução temporal de forma relativamente simples. O estudo mostra como a análise matemática pode contribuir para entender as trajetórias de doenças infecciosas e suas implicações na saúde pública.

Palavras-chave: Epidemiologia. Modelagem Matemática. Gonorreia. COVID-19. Logística.

ABSTRACT

This work presents mathematical models for two diseases that do not confer permanent immunity after infection: gonorrhea and COVID-19. In the case of gonorrhea, nonlinear differential equations are introduced to describe the dynamics between susceptible and infected individuals within a given population. We analyze the stability of equilibrium points and the conditions leading to extinction or endemic persistence.

Subsequently, we address COVID-19 using data analyzed from São Carlos, SP. Instead of employing a classical compartmental model, we resort to a logistic curve fit to describe the accumulated number of cases, discussing how this type of fit can represent the temporal evolution in a relatively simple manner. The study demonstrates how mathematical analysis can contribute to understanding the trajectories of infectious diseases and their implications for public health.

Keywords: Epidemiology. Mathematical Modeling. Gonorrhea. COVID-19. Logistics.

Lista de Figuras

Figura 1 – Item (a) do Teorema.	29
Figura 2 – Item (b) do Teorema.	30
Figura 3 – Número em milhares de casos de gonorreia de 1950 a 1973.	34
Figura 4 – Ajuste Logístico COVID-19	53

Sumário

1	Introdução	9
2	Preliminares	10
2.1	Estabilidade de Sistemas Lineares	10
2.2	Estabilidade de Sistemas Quase-Lineares	14
2.3	Comportamento das Soluções para Tempos Suficientemente Grandes	17
2.4	Métodos Numéricos	19
2.4.1	Formulação Matemática da Otimização	19
3	A disseminação da gonorréia	23
3.1	O Modelo Matemático	23
3.2	Análise global	24
3.3	Análise local	30
3.4	Dados Reais	34
4	Disseminação da Covid-19	36
4.1	Como se dá o contágio	36
4.2	Fundamentação Matemática da Curva Logística	36
4.3	Outro modelo	38
4.4	Aplicação Prática	39
4.4.1	Coleta e Preparação dos Dados	40
4.4.2	Ajuste	43
4.4.3	Comparando o Modelo com Dados Reais	48
4.4.4	Análise Epidemiológica	52
5	Comentários	55
	Referências	56

1. Introdução

No Capítulo 2, Preliminares, abordamos os principais conceitos utilizados no trabalho.

No Capítulo 3, utilizamos resultados como a estabilidade de sistemas lineares e quase-lineares, o Teorema de Poincaré–Bendixson para analisar um sistema de equações diferenciais não lineares que modela a disseminação da gonorreia, o modelo divide a população sexualmente ativa e promíscua em dois grupos (homens e mulheres) e divide cada um entre suscetíveis e infecciosos. São levadas em conta diferenças entre cada um dos grupos, por exemplo, nas taxas de cura e infecção. A análise global do sistema é feita dividindo o domínio em regiões, onde os sinais das derivadas permanecem fixos. Concluimos que, sob certas condições, o modelo prevê tanto a persistência endêmica quanto a eventual extinção da doença.

Em seguida, no Capítulo 4, utilizamos algoritmos de otimização e ajuste introduzidos nas preliminares, que foram todos implementados em Python. Coletamos, filtramos e analisamos os dados disponíveis no Painel Coronavírus do Ministério da Saúde. Com os mesmos, não apenas obtivemos estimativas para os parâmetros K , r e t_0 na equação logística, como também obtivemos o quão bem nosso modelo se adapta à situação real, utilizando o teste de Shapiro–Wilk.

2. Preliminares

Neste capítulo enunciaremos alguns resultados baseados em [Braun \(1993\)](#), [Nocedal e Wright \(2006\)](#) e [Shapiro e Wilk \(1965\)](#) que serão utilizados nesta monografia.

Considere a equação diferencial $\dot{x}(t) = f(t, x)$, onde f é não linear. Infelizmente não se conhecem métodos para obtenção de soluções explícitas dessas equações, no entanto, na maioria das aplicações conseguimos obter propriedades dessas soluções, tais como:

- existem soluções constantes (se existirem são chamadas pontos de equilíbrio)?
- o que ocorrem com as soluções $x(t)$ dessa equação, isto é, elas tendem a valores de equilíbrio ou tendem a uma solução periódica?
- duas soluções com condições iniciais próximas, continuarão assim para todo t ? Esta questão é conhecida como problema de estabilidade.

Responder a estas perguntas é conhecido como teoria qualitativa das equações diferenciais.

A resposta da pergunta dois pode prever o comportamento a longo prazo do sistema estudado. A questão de estabilidade é de primordial importância nos problemas físicos, já que nunca pode se medir com precisão as condições iniciais. É uma questão difícil de resolver já que não conseguimos expressões explícitas das soluções. O único caso que conseguimos resolver é quando temos $\dot{x} = f(x)$, isto é, uma equação autônoma. Mesmo assim, conseguimos resolver somente dois casos: quando $\dot{x} = Ax$ ou quando o interesse é na estabilidade da solução $x = 0$ de $\dot{x} = f(x)$, que é o suficiente para nossos propósitos.

2.1 Estabilidade de Sistemas Lineares

Seja $x = \phi(t)$ uma solução da equação diferencial

$$\dot{x} = f(x). \tag{2.1}$$

Estamos interessados em determinar, particularmente na equação estudada, se uma solução $x = \phi(t)$ do sistema diferencial autônomo é estável ou instável. Isto é, determinar se qualquer outra solução do sistema que começa suficientemente próxima de $\phi(t)$ em $t = 0$ permanece próxima a ela por todo o tempo futuro, $t \geq 0$. A solução $\phi(t)$ é instável se existe pelo menos uma solução $\psi(t)$ que começa perto de $\phi(t)$ em $t = 0$ mas não permanece perto para todo $t \geq 0$. Mais precisamente

Definição 2.1: A solução $\phi(t)$ é estável se para todo $\varepsilon > 0$ existe $\delta = \delta(\varepsilon)$ tal que

$$\|\psi_j(t) - \phi_j(t)\| < \varepsilon \quad \text{se} \quad \|\psi_j(0) - \phi_j(0)\| < \delta(\varepsilon), \quad j = 1, \dots, n$$

para toda solução $\psi(t)$ de (2.1).

O problema da estabilidade pode ser resolvido completamente para todas as soluções da equação diferencial linear

$$\dot{x} = Ax. \quad (2.2)$$

Isto não é surpreendente, pois a equação (2.2) pode ser resolvida exatamente. Temos o seguinte teorema:

Teorema 2.1: (a) Toda solução $x = \phi(t)$ de (2.2) é estável se todos os valores característicos de A têm parte real negativa.

(b) Toda solução $x = \phi(t)$ de (2.2) é instável se há ao menos um valor característico de A que possui parte real positiva.

(c) Suponha que todos os valores característicos de A tenham parte real ≤ 0 e que $\lambda_j = i\omega_j$, com $i = \sqrt{-1}$, tenham parte real igual a zero. Suponha ainda que λ_j seja de multiplicidade k_j . Isso significa que o polinômio característico de A pode ser fatorado como

$$p(\lambda) = (\lambda - i\omega_j)^{k_j} \dots (\lambda - i\omega_j)^{k_q} q(\lambda)$$

onde todas as raízes de $q(\lambda)$ têm parte real negativa. Então, toda solução $x = \phi(t)$ de (2.1) é estável se A tem k_j vetores característicos linearmente independentes para cada valor característico $\lambda_j = i\omega_j$. Caso contrário, todas as soluções $\phi(t)$ são instáveis.

O primeiro passo para demonstrar o teorema 2.1 é mostrar que toda solução $\phi(t)$ é estável (ou instável) se a solução de equilíbrio $x(t) \equiv 0$ o é. Para isso, seja $\psi(t)$ qualquer solução de (2.2). Observe-se que $z(t) = \phi(t) - \psi(t)$ é novamente uma solução de (2.2). Portanto, se a solução de equilíbrio $x(t) \equiv 0$ é estável, então $z(t) = \phi(t) - \psi(t)$ será pequena se $z(0) = \phi(0) - \psi(0)$ for suficientemente pequeno. Assim, toda solução $\phi(t)$ de (2.2) será estável.

Por outro lado, suponha que $x(t) \equiv 0$ seja instável. Então, existe uma solução $x = h(t)$ que é inicialmente muito pequena, mas se torna muito grande quando t tende ao infinito. A função $\psi(t) = \phi(t) + h(t)$ é claramente uma solução de (2.2). Além disso, $\psi(t)$ está inicialmente muito próxima de $\phi(t)$, mas diverge de $\phi(t)$ quando t cresce. Portanto, toda solução $x = \phi(t)$ de (2.2) é instável.

O próximo passo na demonstração do Teorema 2.1 é reduzir o problema de verificar se n quantidades $\psi_j(t)$, $j = 1, \dots, n$, são pequenas ao problema mais simples de mostrar unicamente que a norma de um vetor é pequena.

Definição 2.2: Seja

$$x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

um vetor com n componentes. Os números x_1, \dots, x_n podem ser reais ou complexos. Define-se a *norma* de x , denotada por $\|x\|$, como

$$\|x\| \equiv \max(|x_1|, |x_2|, \dots, |x_n|).$$

.

A demonstração do teorema 2.1 a seguir é válida para qualquer n .

Demonstração: (a) Toda solução $x = \psi(t)$ de $\dot{x} = Ax$ é da forma $\psi(t) = e^{At}\psi(0)$. Seja $\varphi_{ij}(t)$ o elemento j da matriz e^{At} e sejam $\psi_1^0, \dots, \psi_n^0$ as componentes de $\psi(0)$. Então, a componente j de $\psi(t)$ é dada por

$$\psi_j(t) = \sum_{i=1}^n \varphi_{ji}(t)\psi_i^0.$$

Suponha que todos os valores característicos de A tenham parte real negativa. Seja $-\alpha_1$ a maior das partes reais dos valores característicos de A . Seja $\alpha \in \mathbb{R}$ tal que $-\alpha_1 < -\alpha < 0$. Pode-se encontrar um número $K > 0$ tal que $|\varphi_{ji}(t)| \leq Ke^{-\alpha t}$ para todo $t \geq 0$. Assim,

$$|\psi_j(t)| \leq \sum_{i=1}^n Ke^{-\alpha t}|\psi_i^0| = Ke^{-\alpha t} \sum_{i=1}^n |\psi_i^0|.$$

Como $\|\psi(0)\| = \max(|\psi_1^0|, \dots, |\psi_n^0|)$, então

$$\|\psi(t)\| \leq Ke^{-\alpha t}\|\psi(0)\|.$$

Seja $\varepsilon > 0$. Escolha $\delta(\varepsilon) = \varepsilon/(Ke^{-\alpha})$. Então, $\|\psi(t)\| < \varepsilon$ sempre que $\|\psi(0)\| < \delta(\varepsilon)$. Logo,

$$\|\psi(t)\| < \varepsilon \quad \text{se} \quad \|\psi(0)\| < \delta(\varepsilon).$$

Portanto, a solução de equilíbrio $x(t) \equiv 0$ é estável.

(b) Seja λ um valor característico de A com parte real positiva e seja v um vetor característico de A com valor característico λ . Então, $\psi(t) = ce^{\lambda t}v$ é uma solução de $\dot{x} = Ax$ para qualquer constante c . Se λ é real, então v também o é e $\|\psi(t)\| = |c|e^{\lambda t}\|v\|$. Claramente, $\|\psi(t)\|$ tende a infinito quando t tende a infinito, para qualquer escolha de $c \neq 0$, independentemente de quão pequeno seja. Portanto, $x(t) \equiv 0$ é instável.

Se $\lambda = \alpha + i\beta$ for complexo, então $v = v^1 + iv^2$ também é. Nesse caso,

$$\begin{aligned} e^{(\alpha+i\beta)t}(v^1 + iv^2) &= e^{\alpha t}(\cos \beta t + i \operatorname{sen} \beta t)(v^1 + iv^2) \\ &= e^{\alpha t} \left[(v^1 \cos \beta t - v^2 \operatorname{sen} \beta t) + i(v^1 \operatorname{sen} \beta t + v^2 \cos \beta t) \right] \end{aligned}$$

ou seja, uma solução com valores complexos de (2.2). Assim,

$$\psi^1(t) = ce^{\alpha t}(v^1 \cos \beta t - v^2 \operatorname{sen} \beta t)$$

é uma solução com valores reais de (2.2), para qualquer escolha da constante c . Observa-se que $\|\psi^1(t)\|$ não está limitado quando t tende a infinito se c e pelo menos um dos vetores v^1 ou v^2 forem diferentes de zero. Portanto, $x(t) \equiv 0$ é instável.

(c) Se A tem k_j vetores característicos linearmente independentes para cada valor característico $\lambda_j = i\omega_j$ de multiplicidade k_j , então pode-se encontrar uma constante K tal que $|(e^{At})_{ij}| \leq K$. Nesse caso, $\|\psi(t)\| \leq nK\|\psi(0)\|$ para toda solução $\psi(t)$ de (2.2). Da demonstração do item (a), segue imediatamente que $x(t) \equiv 0$ é estável. Por outro lado, se A tem menos de k_j vetores característicos linearmente independentes com valor característico $\lambda_j = i\omega_j$, então $\dot{x} = Ax$ tem soluções $\psi(t)$ da forma

$$\psi(t) = ce^{i\omega_j t} [v + t(A - i\omega_j I)v]$$

onde $(A - i\omega_j I)v \neq 0$. Se $\sigma_j = 0$, então $\psi(t) = c(v + tAv)$ toma valores reais. Além disso, $\|\psi(t)\|$ não está limitado quando t tende a infinito, para qualquer escolha de $c \neq 0$. De maneira similar, tanto a parte real quanto a imaginária de $\psi(t)$ não estão limitadas em norma para $\psi(0) \neq 0$ arbitrariamente pequeno, se $\sigma_j \neq 0$. Portanto, a solução de equilíbrio $x(t) \equiv 0$ é instável.

Se todos os valores característicos de A têm parte real negativa, então toda solução $x(t)$ de $\dot{x} = Ax$ tende a zero quando t tende a infinito. Isso se segue imediatamente da estimativa

$$\|x(t)\| \leq Ke^{-\alpha t} \|x(0)\|$$

obtida na demonstração da parte (a) do teorema 2.1. Assim, não apenas a solução de equilíbrio $x(t) \equiv 0$ é estável, mas toda solução $\psi(t)$ de (2.2) tende a ela quando t tende a infinito. Esse tipo de estabilidade tão forte é conhecido como *estabilidade assintótica*. \square

Definição 2.3: Uma solução $x = \phi(t)$ de (2.1) é assintoticamente estável se for estável e se toda solução $\psi(t)$ que começa suficientemente perto de $\phi(t)$ tender a $\phi(t)$ quando t tende a infinito. Em particular, uma solução de equilíbrio $x(t) = x^0$ de (2.1) é assintoticamente estável se toda solução $x = \psi(t)$ de (2.1) que começa suficientemente próxima de x^0 no instante $t = 0$ não apenas permanecer próxima de x^0 para todo instante posterior, mas também tender a x^0 quando t tende a infinito.

Observação 2.2: A estabilidade assintótica de qualquer solução $x = \phi(t)$ de (2.2) é claramente equivalente à estabilidade assintótica da solução de equilíbrio $x(t) \equiv 0$.

2.2 Estabilidade de Sistemas Quase-Lineares

Consideremos agora o sistema

$$\dot{x} = Ax + g(x) \quad (2.3)$$

onde

$$g(x) = \begin{bmatrix} g_1(x) \\ \vdots \\ g_n(x) \end{bmatrix}$$

é tal que

$$\frac{g_1(x)}{\max(|x_1|, \dots, |x_n|)}, \dots, \frac{g_n(x)}{\max(|x_1|, \dots, |x_n|)}$$

e são funções contínuas de x_1, \dots, x_n que se anulam para $x_1 = \dots = x_n = 0$. Se cada uma das componentes de $g(x)$ for um polinômio em x_1, \dots, x_n que começa com termos de ordem 2 ou superior, então isto ocorre.

Se $g(0) = 0$, então $x(t) \equiv 0$ é uma solução de equilíbrio de (2.3). Parece plausível que a estabilidade da solução de equilíbrio $x(t) \equiv 0$ de (2.3) deva estar relacionada à estabilidade da equação "aproximada" $\dot{x} = Ax$. No próximo teorema, mostra-se que essa ideia é quase correta.

Teorema 2.3: Suponha que a função com valores vetoriais

$$\frac{g(x)}{\|x\|} \equiv \frac{g(x)}{\max(|x_1|, \dots, |x_n|)}$$

seja uma função contínua de x_1, \dots, x_n que se anula para $x = 0$. Então:

- (a) A solução de equilíbrio $x(t) \equiv 0$ de (2.3) é assintoticamente estável se a solução de equilíbrio $x(t) \equiv 0$ da equação "linearizada" $\dot{x} = Ax$ for assintoticamente estável. De maneira equivalente, a solução $x(t) \equiv 0$ de (2.3) é assintoticamente estável se todos os valores característicos de A tiverem parte real negativa.
- (b) A solução de equilíbrio $x(t) \equiv 0$ de (2.3) é instável se ao menos um valor característico de A tiver parte real positiva.
- (c) A estabilidade da solução de equilíbrio $x(t) \equiv 0$ de (2.3) não pode ser determinada a partir da estabilidade da solução de equilíbrio $x(t) \equiv 0$ de $\dot{x} = Ax$ se todos os valores característicos de A tiverem parte real ≤ 0 , mas ao menos um valor característico de A tiver parte real igual a zero.

Demonstraremos aqui apenas o item (a).

Demonstração: Pela fórmula da variação de parâmetros, toda solução $x(t)$ de (2.3) pode ser escrita na forma

$$x(t) = e^{At}x(0) + \int_0^t e^{A(t-s)}g(x(s))ds. \quad (2.4)$$

Deseja-se mostrar que $\|x(t)\|$ tende a zero quando t tende a infinito. Para isso, recorde-se que, se todos os valores característicos ou autovalores de A têm parte real negativa, então é possível encontrar constantes K e α tais que

$$\|e^{At}x(0)\| \leq Ke^{-\alpha t}\|x(0)\|$$

e

$$\|e^{A(t-s)}g(x(s))\| \leq Ke^{-\alpha(t-s)}\|g(x(s))\|.$$

Além disso, é possível encontrar uma constante positiva σ tal que

$$\|g(x)\| < \frac{\alpha}{2K}\|x\| \quad \text{se } \|x\| < \sigma.$$

Isso se segue imediatamente da suposição de que $g(x)/\|x\|$ é contínua e se anula em $x = 0$. Portanto, a equação (2.4) implica que

$$\begin{aligned} \|x(t)\| &\leq \|e^{At}x(0)\| + \int_0^t \|e^{A(t-s)}g(x(s))\|ds \\ &< Ke^{-\alpha t}\|x(0)\| + \frac{\alpha}{2} \int_0^t e^{-\alpha(t-s)}\|x(s)\|ds. \end{aligned}$$

Multiplicando por $e^{\alpha t}$ ambos os lados da desigualdade, obtém-se

$$e^{\alpha t}\|x(t)\| \leq K\|x(0)\| + \frac{\alpha}{2} \int_0^t e^{\alpha s}\|x(s)\|ds. \quad (2.5)$$

A desigualdade (2.5) pode ser simplificada definindo $z(t) = e^{\alpha t}\|x(t)\|$, de modo que

$$z(t) < K\|x(0)\| + \frac{\alpha}{2} \int_0^t z(s)ds. \quad (2.6)$$

Definindo

$$U(t) = \frac{\alpha}{2} \int_0^t z(s)ds,$$

então,

$$\frac{dU(t)}{dt} = \frac{\alpha}{2}z(t) < \frac{\alpha}{2}K\|x(0)\| + \frac{\alpha}{2}U(t),$$

ou equivalentemente,

$$\frac{dU(t)}{dt} - \frac{\alpha}{2}U(t) < \frac{\alpha K}{2}\|x(0)\|.$$

Multiplicando ambos os lados da desigualdade pelo fator de integração $e^{-\alpha t/2}$, obtém-se

$$\frac{d}{dt}e^{-\alpha t/2}U < \frac{\alpha K}{2}\|x(0)\|e^{-\alpha t/2},$$

o que implica que

$$\frac{d}{dt}e^{-\alpha t/2}[U(t) + K\|x(0)\|] < 0.$$

Portanto,

$$e^{-\alpha t/2}[U(t) + K\|x(0)\|] < U(0) + K\|x(0)\| = K\|x(0)\|,$$

de modo que

$$U(t) \leq -K\|x(0)\| + K\|x(0)\|e^{\alpha t/2}.$$

Retornando agora à desigualdade (2.6), temos

$$\begin{aligned}\|x(t)\| &= e^{-\alpha t/2}z(t) \leq e^{-\alpha t}[K\|x(0)\| + U(t)] \\ &\leq K\|x(0)\|e^{-\alpha t/2}.\end{aligned}\tag{2.7}$$

Agora, se $\|x(0)\| \leq \sigma/K$, então a desigualdade (2.7) garante que $\|x(t)\| \leq \sigma$ para todo instante posterior t . Portanto, a desigualdade (2.7) é válida para todo $t \geq 0$ se $\|x(0)\| \leq \sigma/K$. Finalmente, observa-se da desigualdade (2.7) que $\|x(t)\| \leq K\|x(0)\|$ e que $x(t)$ tende a zero quando t tende a infinito. Assim, a solução de equilíbrio $x(t) \equiv 0$ de (2.3) é assintoticamente estável. \square

O Teorema 2.3 serve também para determinar a estabilidade das soluções de equilíbrio de equações diferenciais autônomas arbitrárias. Seja x^0 um valor de equilíbrio da equação diferencial

$$\dot{x} = f(x)\tag{2.8}$$

e defina-se $z(t) = x(t) - x^0$. Então,

$$\dot{z} = \dot{x} = f(x^0 + z).\tag{2.9}$$

Obviamente, $z(t) \equiv 0$ é uma solução de equilíbrio de (2.9) e a estabilidade de $x(t) \equiv x^0$ é equivalente à estabilidade de $z(t) \equiv 0$.

Vamos mostrar que $f(x^0 + z) = Az + g(z)$, onde $g(z)$ é pequeno em comparação com z .

Lema 2.1: Suponha que $f(x)$ tem duas derivadas parciais contínuas com respeito a cada uma de suas variáveis x_1, \dots, x_n . Então, $f(x^0 + z)$ pode ser escrita na forma

$$f(x^0 + z) = f(x^0) + Az + g(z)\tag{2.10}$$

onde $g(z)/\max(|z_1|, \dots, |z_n|)$ é uma função contínua de z que se anula para $z = 0$.

Demonstração: A equação (2.10) é uma consequência imediata do Teorema de Taylor, que

estabelece que cada uma das componentes $f_j(x^0 + z)$ de $f(x^0 + z)$ pode ser escrita na forma

$$f_j(x^0 + z) = f_j(x^0) + \frac{\partial f_j(x^0)}{\partial x_1} z_1 + \cdots + \frac{\partial f_j(x^0)}{\partial x_n} z_n + g_j(z)$$

onde $g_j(z)/\max(|z_1|, \dots, |z_n|)$ é uma função contínua de z que se anula para $z = 0$. Portanto,

$$f(x^0 + z) = f(x^0) + Az + g(z)$$

onde

$$A = \begin{bmatrix} \frac{\partial f_1(x^0)}{\partial x_1} & \cdots & \frac{\partial f_1(x^0)}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial f_n(x^0)}{\partial x_1} & \cdots & \frac{\partial f_n(x^0)}{\partial x_n} \end{bmatrix}.$$

□

O Teorema 2.3 e o Lema 2.1 fornecem o seguinte método para determinar se uma solução de equilíbrio $x(t) \equiv x^0$ de $\dot{x} = f(x)$ é estável ou instável:

1. Expressar $z = x - x^0$.
2. Escrever $f(x^0 + z)$ na forma $Az + g(z)$, onde $g(z)$ é um polinômio em z_1, \dots, z_n com valores vetoriais, que começa com termos de ordem dois ou maior.
3. Calcular os valores característicos de A . Se todos os autovalores de A tiverem parte real negativa, então $x(t) \equiv x^0$ é assintoticamente estável. Se algum valor característico de A tiver parte real positiva, então $x(t) \equiv x^0$ é instável.

2.3 Comportamento das Soluções para Tempos Suficientemente Grandes

Nessa seção vemos um resultado que determina o comportamento para valores grandes do tempo de todas as soluções da equação diferencial

$$\dot{x} = f(x) \tag{2.11}$$

Para $f(x) = Ax$, temos os seguintes tipos de comportamento:

- 1. $x(t)$ é constante como função do tempo;
- 2. $x(t)$ é uma função periódica do tempo;
- 3. $x(t)$ não é limitada quando $t \rightarrow \infty$;
- 4. $x(t)$ tende a um ponto de equilíbrio quando $t \rightarrow \infty$.

Na seção anterior, já vimos uma solução parcial no caso $f(x)$ não linear. Abaixo, demonstramos lemas que desempenham um papel extremamente importante.

Lema 2.2: Seja $g(t)$ uma função monótona crescente (decrecente) do tempo para $t \geq t_0$, com $g(t) \leq c$ (ou $\geq c$) para alguma constante c . Então, $g(t)$ tem um limite quando t tende a infinito.

Demonstração: Suponha que $g(t)$ é monótona crescente para $t \geq t_0$ e que $g(t)$ está limitada superiormente. Seja l a mínima cota superior de g , ou seja, o menor número não excedido pelos valores de $g(t)$ para $t \geq t_0$. Esse número deve ser o limite de $g(t)$ quando t tende a infinito. Para demonstrar isso, tome $\varepsilon > 0$ e observe que existe um tempo $t_\varepsilon \geq t_0$ tal que $l - g(t_\varepsilon) < \varepsilon$. Se tal tempo t_ε não existisse, então l não seria a mínima cota superior de g . Como $g(t)$ é monótona, tem-se que $l - g(t) < \varepsilon$ para $t \geq t_\varepsilon$. Isso mostra que $l = \lim_{t \rightarrow \infty} g(t)$. \square

Lema 2.3: Suponha que uma solução $x(t)$ de (2.11) tende a um vetor ξ quando t tende a infinito. Então, ξ é um ponto de equilíbrio de (2.11).

Demonstração: Suponha que $x(t)$ tende a ξ quando t tende a infinito. Então, $x_j(t)$ tende a ξ_j , onde ξ_j é a componente j de ξ . Isso implica que $|x_j(t_1) - x_j(t_2)|$ tende a zero quando t_1 e t_2 tendem a infinito, pois

$$\begin{aligned} |x_j(t_1) - x_j(t_2)| &= |(x_j(t_1) - \xi_j) + (\xi_j - x_j(t_2))| \\ &\leq |x_j(t_1) - \xi_j| + |x_j(t_2) - \xi_j|. \end{aligned}$$

Em particular, seja $t_1 = t$ e $t_2 = t + h$ para algum número positivo fixo h . Então, $|x_j(t+h) - x_j(t)|$ tende a zero quando t tende a infinito. No entanto,

$$x_j(t+h) - x_j(t) = h \frac{dx_j(\tau)}{dt} = h f_j(x_1(\tau), \dots, x_n(\tau)),$$

onde τ é algum número entre t e $t+h$. Observe-se, por fim, que $f_j(x_1(\tau), \dots, x_n(\tau))$ deve tender a $f_j(\xi_1, \dots, \xi_n)$ quando t tende a infinito. Portanto, $f_j(\xi_1, \dots, \xi_n) = 0$, para $j = 1, 2, \dots, n$, concluindo assim a demonstração do Lema 1. \square

Teorema 2.4 (Teorema de Poincaré-Bendixson): Suponha que uma solução $x = x(t)$, $y = y(t)$ do sistema de equações diferenciais, $\dot{x} = f(x, y)$, $\dot{y} = g(x, y)$, permaneça em uma região limitada do plano que não contém pontos de equilíbrio do sistema. Então sua órbita deve espiralar para uma curva fechada simples a qual é a órbita de uma solução periódica do sistema.

O próximo lema é um exercício proposto em Braun (1993) (Exercício 11 - página 437).

Lema 2.4: Suponhamos que $x(t)$, $y(t)$ seja uma solução periódica do sistema $\dot{x} = f(x, y)$, $\dot{y} = g(x, y)$ e C a órbita dessa solução. Então a integral de linha

$$\int_C f(x, y) dy - g(x, y) dx$$

para essa curva C é zero. Também o sistema não pode ter uma solução periódica inteiramente contida numa região simplesmente conexa D do plano Oxy onde $f_x + g_y$ tenha o mesmo sinal.

De fato, pois tomando uma parametrização $\gamma(t) = (x(t), y(t))$, $a \leq t \leq b$, temos

$$\int_C f(x, y)dy - g(x, y)dx = \int_\gamma [f(x(t), y(t))\dot{y}(t) - g(x(t), y(t))\dot{x}(t)]dt = \int_a^b [\dot{x}y - y\dot{x}]dt = 0.$$

Para provar a segunda afirmação do lema, suponha que o sistema tenha uma solução periódica inteiramente contida numa região simplesmente conexa do plano Oxy onde $f_x + g_y$ tem o mesmo sinal. Pelo que vimos e pelo Teorema de Green (para R região delimitada por C) teríamos

$$0 = \int_C f(x, y)dy + g(x, y)dx = \int \int_R [f_x(x, y) + g_y(x, y)]dxdy,$$

o que implica, $f_x(x, y) + g_y(x, y) = 0$, uma contradição.

2.4 Métodos Numéricos

Aqui, traremos uma breve introdução aos métodos numéricos utilizados no capítulo 4, seção 4.4.2 (mais precisamente, os resultados 2.4, 2.5 e 2.6) para otimização e ajuste.

2.4.1 Formulação Matemática da Otimização

Algoritmos de otimização são iterativos. Eles começam com um palpite inicial para a variável x e geram uma sequência de estimativas melhoradas (chamadas “iterações”) até que terminem, de preferência em uma solução. A estratégia usada para passar de uma iteração para a seguinte distingue um algoritmo de outro. A maioria das estratégias faz uso dos valores da função objetivo f , das funções de restrição c_i e, possivelmente, das primeiras e segundas derivadas dessas funções. Alguns algoritmos acumulam informações coletadas em iterações anteriores, enquanto outros usam apenas informações locais obtidas no ponto atual. Independentemente dessas particularidades, bons algoritmos devem possuir as seguintes propriedades:

- **Robustez.** Devem ter bom desempenho em uma ampla variedade de problemas em sua classe, para todos os valores razoáveis do ponto inicial.
- **Eficiência.** Não devem exigir tempo ou armazenamento computacional excessivo.
- **Precisão.** Devem ser capazes de identificar uma solução com exatidão, sem serem excessivamente sensíveis a erros nos dados ou a erros de arredondamento aritmético que ocorrem quando o algoritmo é implementado em um computador.

Matematicamente, a otimização é a minimização ou maximização de uma função sujeita a restrições em suas variáveis. Usamos a seguinte notação:

- x é o vetor de variáveis, também chamado de incógnitas ou parâmetros;

- f é a função objetivo, uma função (escalar) de x que desejamos maximizar ou minimizar;
- c_i são funções de restrição, que são funções escalares de x que definem certas equações e desigualdades que o vetor desconhecido x deve satisfazer.

Iniciando em x_0 , algoritmos de otimização geram uma sequência de iterações $\{x_k\}_{k=0}^{\infty}$ que terminam quando não é mais possível fazer progresso ou quando parece que um ponto de solução foi aproximado com precisão suficiente. Ao decidir como avançar de uma iteração x_k para a próxima, os algoritmos usam informações sobre a função f em x_k e, possivelmente, também informações de iterações anteriores x_0, x_1, \dots, x_{k-1} . Eles utilizam essas informações para produzir uma nova iteração x_{k+1} com valor de função menor que x_k . (Existem algoritmos não monótonos que não exigem uma diminuição em f a cada passo, mas mesmo esses algoritmos requerem que f seja reduzida após um número prescrito m de iterações, isto é, $f(x_t) < f(x_{t-m})$.)

As informações coletadas sobre f são usadas para construir uma função modelo m_k , cujo comportamento próximo ao ponto atual x_k é semelhante ao de f . Como o modelo m_k pode não representar com precisão a função f quando x está muito distante de x_k , restringimos a busca pelo minimizador de m_k a uma região ao redor de x_k . Em outras palavras, encontra-se o passo candidato p ao resolver aproximadamente o seguinte subproblema:

$$\min_p m_k(x_k + p), \quad (2.12)$$

onde $x_k + p$ permanece dentro da região de confiança.

Se a solução candidata não produzir uma diminuição suficiente em f , concluímos que a região de confiança é grande demais, reduzimos seu tamanho e resolvemos novamente (2.12). Geralmente, a região de confiança é uma bola definida por $\|p\| \leq \Delta$, onde o escalar $\Delta > 0$ é chamado de raio da região de confiança. Regiões de confiança elípticas ou em formato de caixa também podem ser usadas.

O modelo m_k em (2.12) costuma ser definido como uma função quadrática da forma

$$m_k(x_k + p) = f_k + p^T \nabla f_k + \frac{1}{2} p^T B_k p,$$

onde f_k , ∇f_k e B_k são, respectivamente, um escalar, um vetor e uma matriz. Conforme a notação indica, f_k e ∇f_k são escolhidos como os valores da função e do gradiente em x_k , de modo que m_k e f coincidam na primeira ordem na iteração atual x_k . A matriz B_k é o Hessiano $\nabla^2 f$ ou alguma aproximação dele.

Definição 2.4 (Soma dos Quadrados dos Resíduos (SSR)): Sejam $\{(t_i, C_{\text{obs}}(t_i))\}_{i=1}^n$ os dados experimentais, onde:

- t_i representa os instantes de tempo em que foram registradas as contagens de casos acumulados;
- $C_{\text{obs}}(t_i)$ é o número observado de casos acumulados no instante t_i .

Definindo o modelo logístico

$$C(t; K, r, t_0) = \frac{K}{1 + e^{-r(t-t_0)}},$$

a soma dos quadrados dos resíduos (SSR), é dada por

$$\text{SSR}(K, r, t_0) = \sum_{i=1}^n \left[C_{\text{obs}}(t_i) - \frac{K}{1 + e^{-r(t_i-t_0)}} \right]^2.$$

A minimização da SSR em relação aos parâmetros K , r e t_0 permite obter os valores que proporcionam o melhor ajuste do modelo aos dados observados.

Definição 2.5 (Algoritmo de Levenberg–Marquardt): Seja

$$\Phi(\mathbf{x}) = \frac{1}{2} \|F(\mathbf{x})\|^2,$$

a função associada a um problema de mínimos quadrados não lineares, onde $F : \mathbb{R}^n \rightarrow \mathbb{R}^m$ é uma função diferenciável e $\mathbf{x} \in \mathbb{R}^n$ representa o vetor de parâmetros. O algoritmo de Levenberg–Marquardt gera uma sequência $\{\mathbf{x}_k\}$, com a iterada $\mathbf{x}_{k+1} = \mathbf{x}_k + \delta$ com δ satisfazendo

$$\left[J(\mathbf{x}_k)^T J(\mathbf{x}_k) + \lambda I \right] \delta = -J(\mathbf{x}_k)^T F(\mathbf{x}_k),$$

onde

- $J(\mathbf{x}_k)$ é a matriz jacobiana de F em \mathbf{x}_k ;
- $\lambda > 0$ é um parâmetro de amortecimento;
- I é a matriz identidade.

Sob condições regulares, a sequência $\{\mathbf{x}_k\}$ converge localmente para um minimizador de $\Phi(\mathbf{x})$.

A inclusão do termo λI melhora a condição do sistema linear quando $J(\mathbf{x}_k)^T J(\mathbf{x}_k)$ é singular ou mal condicionado, possibilitando uma transição entre o método de Gauss–Newton (quando λ é pequeno) e o gradiente descendente (quando λ é grande). Dessa forma, o método se mostra mais robusto e converge localmente sob hipóteses regulares.

Definição 2.6 (Teste de Shapiro–Wilk): Seja x_1, x_2, \dots, x_n uma amostra de dados analisados com média amostral \bar{x} e ordenados em ordem crescente:

$$x_1 \leq x_2 \leq \dots \leq x_n.$$

O teste de Shapiro–Wilk utiliza a estatística

$$W = \frac{(\sum_{i=1}^n a_i x_i)^2}{\sum_{i=1}^n (x_i - \bar{x})^2},$$

na qual os coeficientes a_i são determinados a partir dos quantis esperados e das covariâncias dos valores ordenados de uma distribuição normal padrão. Se os dados são provenientes de uma distribuição normal, o valor de W tende a ser próximo de 1; valores significativamente menores indicam desvios da normalidade.

3. A disseminação da gonorréia

Neste capítulo se constrói um modelo matemático para a disseminação da gonorréia. O trabalho se simplifica fortemente pelo fato do período de incubação ser muito curto, 3 a 7 dias, comparado o período de infectividade ativa com outro de longo período de atividade infecciosa. Sendo assim, vamos deduzir um modelo matemático onde vamos supor que os indivíduos se tornam infecciosos imediatamente depois de contrair a doença. Além disso, a gonorreia não confere imunidade a quem a contrai. Sendo assim, temos que considerar que um indivíduo volta a ser suscetível imediatamente após sua recuperação.

3.1 O Modelo Matemático

A população sexualmente ativa e promíscua pode ser dividida em dois grupos: dos suscetíveis e dos infecciosos. Sejam, no instante t , $c_1(t)$ o número de promíscuos do sexo masculino, $c_2(t)$ o número de promíscuos do sexo feminino, $x(t)$ o número de infecciosos do sexo masculino e $y(t)$ o número de infecciosos do sexo feminino. Então o número de homens suscetíveis e mulheres suscetíveis são $c_1(t) - x(t)$ e $c_2(t) - y(t)$, respectivamente. A propagação da gonorréia é presumida ser governada pelas seguintes regras:

- a) Homens infecciosos são curados numa taxa a_1 proporcional ao seu número total, e as mulheres infecciosas são curadas numa taxa a_2 proporcional ao seu número total. A constante a_1 é maior que a constante a_2 já que os homens infectados desenvolvem rapidamente sintomas dolorosos e, portanto, procura assistência médica imediatamente. Mulheres infectadas, por outro lado, geralmente são assintomáticas e, portanto, são infecciosas por um período muito maior.
- b) Novos infectados são adicionados à população masculina numa taxa b_1 proporcional ao número total de homens suscetíveis e mulheres infecciosas. Analogamente, novas infectadas são adicionadas à população feminina numa taxa b_2 proporcional ao número total de mulheres suscetíveis e homens infecciosos.
- c) O número total de homens promíscuos e mulheres promíscuas permanecem em níveis constantes c_1 e c_2 , respectivamente.

Segue imediatamente dessas regras que

$$\begin{aligned} \dot{x}(t) &= -a_1x + b_1(c_1 - x)y \\ \dot{y}(t) &= -a_2y + b_2(c_2 - y)x. \end{aligned} \tag{3.1}$$

Observamos que este sistema trata somente daqueles casos de gonorréia que aparecem de contatos heterossexuais.

3.2 Análise global

Nosso primeiro passo na análise do sistema de equações diferenciais acima é mostrar que é realístico. Especificamente, devemos mostrar que $x(t)$ e $y(t)$ nunca se tornam negativos e que nunca excedem c_1 e c_2 , respectivamente. Isso é mostrado nos Lemas 1 e 2 abaixo.

Lema 3.1: Se $x(t_0)$ e $y(t_0)$ são positivos, então $x(t)$ e $y(t)$ são positivos para todo $t \geq t_0$.

Demonstração: Suponhamos que o lema seja falso. Seja $t^* > t_0$ o primeiro tempo no qual x ou y é zero. Vamos assumir que x se anula primeiro. Então substituindo $t = t^*$ na primeira equação do sistema acima obtemos $\dot{x}(t^*) = b_1 c_1 y(t^*)$. Esse número é positivo (observe que $y(t^*)$ não pode ser zero já que $x = 0$ e $y = 0$ é uma solução de equilíbrio do sistema). Logo $x(t)$ é menor que zero para $t < t^*$ e próximo de t^* . Mas isso contradiz a hipótese de que t^* é o primeiro t no qual $x(t) = 0$. Chegamos à mesma contradição se $y(t^*) = 0$. Assim, ambos $x(t)$ e $y(t)$ são positivos para $t \geq t_0$. \square

Lema 3.2: Se $x(t_0) < c_1$ e $y(t_0) < c_2$ então $x(t) < c_1$ e $y(t) < c_2$ para todo $t \geq t_0$.

Demonstração: Suponhamos que o lema seja falso. Seja $t^* > t_0$ o primeiro tempo no qual $x = c_1$ ou $y = c_2$. Vamos supor que $x(t^*) = c_1$. Substituindo $t = t^*$ na primeira equação do sistema acima obtemos $\dot{x}(t^*) = -a_1 c_1$. Esse número é negativo (observe que $y(t^*)$ não pode ser zero já que pelo lema anterior $x(t)$ e $y(t)$ são positivos para todo $t \geq 0$). Logo $x(t)$ é maior que c_1 para $t < t^*$ e próximo de t^* . Mas isso contradiz a hipótese de que t^* é o primeiro t no qual $x(t) = c_1$. Chegamos à mesma contradição se $y(t^*) = c_2$. Assim, ambos $x(t) < c_1$ e $y(t) < c_2$ para $t \geq t_0$. \square

Tendo mostrado que o sistema é um modelo realístico da gonorréia, veremos que previsões ele faz com relação ao curso dessa doença, isto é, se a taxa de cura for menor que a taxa de contágio numa determinada proporção, então ela tende à um valor fixo não nulo. Caso contrário, a doença tende à extinção. Isto é o que diz o seguinte teorema de epidemiologia.

Teorema 3.1: a) Suponha que $a_1 a_2 < b_1 b_2 c_1 c_2$. Então toda solução $(x(t), y(t))$ do sistema com $0 < x(t_0) < c_1$ e $0 < y(t_0) < c_2$ se aproxima da solução de equilíbrio

$$\bar{x} = \frac{b_1 b_2 c_1 c_2 - a_1 a_2}{a_1 b_2 + b_1 b_2 c_2}, \quad \bar{y} = \frac{b_1 b_2 c_1 c_2 - a_1 a_2}{a_2 b_1 + b_1 b_2 c_1},$$

quando $t \rightarrow \infty$. Em outras palavras, o número total de homens infectados e mulheres infectadas tendem a valores fixados.

b) Suponha que $a_1 a_2 > b_1 b_2 c_1 c_2$. Então toda solução $x(t), y(t)$ do sistema com $0 < x(t_0) < c_1$ e $0 < y(t_0) < c_2$ se aproxima de zero quando $t \rightarrow \infty$. Em outras palavras, a gonorréia finalmente se extingue.

Nosso primeiro passo para provar o item (a) do teorema é dividir o retângulo $0 < x < c_1$, $0 < y < c_2$ em regiões nas quais ambos, \dot{x} e \dot{y} têm sinais fixos. Isto é feito da seguinte maneira. Fazendo $\dot{x} = 0$ no sistema e resolvendo y como função de x temos

$$y = \frac{a_1 x}{b_1(c_1 - x)} \equiv \phi_1(x).$$

Analogamente, fazendo $\dot{y} = 0$ no sistema temos

$$x = \frac{a_2 y}{b_2(c_2 - y)} \text{ ou } y = \frac{b_2 c_2 x}{a_2 + b_2 x} \equiv \phi_2(x).$$

Observe que

- $\phi_1(x)$ e $\phi_2(x)$ são funções monótonas crescentes de x pois

$$\phi_1'(x) = \frac{a_1 b_1 (c_1 - x) + a_1 b_1 x}{(b_1(c_1 - x))^2} = \frac{a_1 b_1 c_1}{(b_1(c_1 - x))^2} > 0$$

e

$$\phi_2'(x) = \frac{b_2 c_2 (a_2 + b_2 x) - b_2^2 c_2 x}{(a_2 + b_2 x)^2} = \frac{a_2 b_2 c_2}{(a_2 + b_2 x)^2} > 0$$

- $\lim_{x \rightarrow c_1} \phi_1(x) = +\infty$ e $\lim_{x \rightarrow +\infty} \phi_2(x) = c_2$
- as curvas $\phi_1(x)$ e $\phi_2(x)$ se interceptam nos pontos de equilíbrio $(0, 0)$ e (\bar{x}, \bar{y}) onde

$$\bar{x} = \frac{b_1 b_2 c_1 c_2 - a_1 a_2}{a_1 b_2 + b_1 b_2 c_2}, \quad \bar{y} = \frac{b_1 b_2 c_1 c_2 - a_1 a_2}{a_2 b_1 + b_1 b_2 c_1}.$$

- $\phi_2(x)$ cresce mais rapidamente que $\phi_1(x)$ em $x = 0$ pois

$$\phi_2'(0) = \frac{b_2 c_2}{a_2} > \frac{a_1}{b_1 c_1} = \phi_1'(0)$$

já que $a_1 a_2 < b_1 b_2 c_1 c_2$. Logo, $\phi_2(x)$ está acima de $\phi_1(x)$ para $0 < x < \bar{x}$ e abaixo para $\bar{x} < x < c_1$, veja Figura 2.

- $\dot{x}(t) > 0$ para todo (x, y) acima da curva $y = \phi_1(x)$ e negativa em qualquer ponto (x, y) abaixo dessa curva. Analogamente, $\dot{y}(t) > 0$ para todo (x, y) abaixo da curva $y = \phi_2(x)$ e negativa em qualquer ponto (x, y) acima dessa curva. Assim, as curvas $y = \phi_1(x)$ e $y = \phi_2(x)$ dividem o retângulo $0 < x < c_1$, $0 < y < c_2$ em quatro regiões nas quais \dot{x} e \dot{y} têm sinais fixos.

Também precisaremos dos seguintes lemas para a prova do item (a) do teorema.

Lema 3.3: Toda solução $(x(t), y(t))$ do Sistema (3.1) que começa na região I em $t = t_0$, permanecerá nesta região e tenderá para a solução de equilíbrio (\bar{x}, \bar{y}) quando $t \rightarrow \infty$.

Demonstração: Suponhamos que uma solução $(x(t), y(t))$ do Sistema (3.1), com $(x(t_0), y(t_0))$ na região I, e que saia dessa região em $t = t^*$. Então $\dot{x}(t^*) = 0$ ou $\dot{y}(t^*) = 0$ já que o único modo da solução deixar a região I é cruzando as curvas $\phi_1(x)$ ou $\phi_2(x)$. Vamos supor $\dot{x}(t^*) = 0$. Derivando ambos os lados da primeira equação do Sistema (3.1) em relação a t e fazendo $t = t^*$ obtemos

$$\ddot{x}(t^*) = b_1(c_1 - x(t^*))\dot{y}(t^*)$$

que é positivo, pois $x(t^*) < c_1$ e $\dot{y}(t^*) > 0$ na curva $y = \phi_1(x)$, $0 < x < \bar{x}$. Logo t^* é um ponto de mínimo de $x(t)$, o que é impossível já que na região I $x(t)$ é sempre crescente.

Analogamente, se $\dot{y}(t^*) = 0$, então

$$\ddot{y}(t^*) = b_2(c_2 - y(t^*))\dot{x}(t^*)$$

é positivo, pois $y(t^*) < c_2$ e $\dot{x}(t^*) > 0$ na curva $y = \phi_2(x)$, $0 < x < \bar{x}$. Logo t^* é um ponto de mínimo de $y(t)$, o que é impossível já que na região I $y(t)$ é sempre crescente.

Concluimos então que toda solução $(x(t), y(t))$ de (3.1) que começa na região I em $t = t_0$ permanecerá nessa região para $t \geq t_0$. Isso implica que,

- $x(t)$ é monótona crescente para $t \geq t_0$ com $x(t) < \bar{x}$ e
- $y(t)$ é monótona crescente para $t \geq 0$ com $y(t) < \bar{y}$.

Pelo Lema 2.2, $\lim_{t \rightarrow \infty} (x(t), y(t)) = (\xi_1, \xi_2)$ e, pelo Lema 2.3, (ξ_1, ξ_2) é um ponto de equilíbrio do Sistema (3.1). Como (ξ_1, ξ_2) não pode ser $(0, 0)$ (pois $x(t)$ e $y(t)$ são crescentes) concluimos que $(\xi_1, \xi_2) = (\bar{x}, \bar{y})$, o que prova o lema. \square

Lema 3.4: Toda solução $(x(t), y(t))$ do Sistema (3.1) que começa na região III em $t = t_0$, permanecerá nesta região e tenderá para a solução de equilíbrio (\bar{x}, \bar{y}) quando $t \rightarrow \infty$.

Demonstração: Suponhamos que uma solução $(x(t), y(t))$ do Sistema (3.1), com $(x(t_0), y(t_0))$ na região III, e que saia dessa região em $t = t^*$. Então $\dot{x}(t^*) = 0$ ou $\dot{y}(t^*) = 0$ já que o único modo da solução deixar a região III é cruzando as curvas $\phi_1(x)$ ou $\phi_2(x)$. Vamos supor $\dot{x}(t^*) = 0$. Derivando ambos os lados da primeira equação do Sistema (3.1) em relação a t e fazendo $t = t^*$ obtemos

$$\ddot{x}(t^*) = b_1(c_1 - x(t^*))\dot{y}(t^*)$$

que é negativo, pois $x(t^*) < c_1$ e $\dot{y}(t^*) < 0$ na curva $\phi_1(x)$, $\bar{x} < x < c_1$. Logo t^* é um ponto de máximo de $x(t)$, o que é impossível já que na região III $x(t)$ é sempre decrescente.

Analogamente, se $\dot{y}(t^*) = 0$, então

$$\ddot{y}(t^*) = b_2(c_2 - y(t^*))\dot{x}(t^*)$$

é negativo, pois $y(t^*) < c_2$ e $\dot{x}(t^*) < 0$ na curva $\phi_2(x)$, $\bar{x} < x < c_1$. Logo t^* é um ponto de máximo de $y(t)$, o que é impossível já que na região III $y(t)$ é sempre decrescente.

Concluimos então que toda solução $(x(t), y(t))$ de (3.1) que começa na região III em $t = t_0$ permanecerá nessa região para $t \geq t_0$. Isso implica que,

- $x(t)$ é monótona decrescente para $t \geq t_0$ com $x(t) > \bar{x}$ e
- $y(t)$ é monótona decrescente para $t \geq 0$ com $y(t) > \bar{y}$.

Pelo Lema 2.2, $\lim_{t \rightarrow \infty} (x(t), y(t)) = (\xi_1, \xi_2)$ e, pelo Lema 2.3, (ξ_1, ξ_2) é um ponto de equilíbrio do Sistema (3.1). Como (ξ_1, ξ_2) não pode ser $(0, 0)$ concluimos que $(\xi_1, \xi_2) = (\bar{x}, \bar{y})$, o que prova o lema. \square

Lema 3.5: Toda solução $(x(t), y(t))$ do Sistema (3.1) que começa na região II em $t = t_0$ e permanece nesta região para $t \geq t_0$ tende para o ponto de equilíbrio (\bar{x}, \bar{y}) quando $t \rightarrow \infty$.

Demonstração: Suponhamos que exista uma solução $(x(t), y(t))$ de (3.1) que permanece na região II para $t \geq t_0$. Então $x(t)$ é monótona decrescente e $y(t)$ é monótona crescente para $t \geq t_0$. Além disso $x(t)$ é maior que zero e $y(t)$ é menor que c_2 , para $t \geq t_0$. Consequentemente, ambas $x(t)$ e $y(t)$ têm limite pelo Lema 2.2, $\lim_{t \rightarrow \infty} x(t) = \xi_1$ e $\lim_{t \rightarrow \infty} y(t) = \xi_2$. Pelo Lema 2.3, o ponto (ξ_1, ξ_2) é de equilíbrio. Mas (ξ_1, ξ_2) não pode ser $(0, 0)$ já que $y(t)$ é crescente na região II para $t \geq t_0$. Portanto concluimos que $(\xi_1, \xi_2) = (\bar{x}, \bar{y})$, o que prova o lema. \square

Lema 3.6: Toda solução $(x(t), y(t))$ do Sistema (3.1) que começa na região IV em $t = t_0$ e permanece nesta região para $t \geq t_0$ tende para o ponto de equilíbrio (\bar{x}, \bar{y}) quando $t \rightarrow \infty$.

Demonstração: Suponhamos que exista uma solução $(x(t), y(t))$ de (3.1) que permaneça na região IV para $t \geq t_0$. Então $x(t)$ é monótona crescente e $y(t)$ é monótona decrescente para $t \geq t_0$. Além disso $x(t)$ é menor que c_1 e $y(t)$ é maior que zero, para $t \geq t_0$. Consequentemente, ambas $x(t)$ e $y(t)$ têm limite pelo Lema 2.2, $\lim_{t \rightarrow \infty} x(t) = \xi_1$ e $\lim_{t \rightarrow \infty} y(t) = \xi_2$. Pelo Lema 2.3, o ponto (ξ_1, ξ_2) é de equilíbrio. Mas (ξ_1, ξ_2) não pode ser $(0, 0)$ já que $x(t)$ é crescente na região IV para $t \geq t_0$. Portanto concluimos que $(\xi_1, \xi_2) = (\bar{x}, \bar{y})$, o que prova o lema. \square

Vamos agora demonstrar o item (a) do teorema 3.1, veja Figura 1.

Demonstração de 3.1 (a): Os Lemas 3.3 e 3.4 estabelecem que qualquer solução $(x(t), y(t))$ do Sistema (3.1) que começa na região I ou III em $t = t_0$ tende para a solução de equilíbrio (\bar{x}, \bar{y}) quando $t \rightarrow \infty$. Analogamente, os Lemas 3.5 e 3.6 estabelecem que qualquer solução $(x(t), y(t))$ do Sistema (3.1) que começa na região II ou IV em $t = t_0$ e que permanece nessas regiões para $t \geq t_0$ deve tender também para a solução de equilíbrio (\bar{x}, \bar{y}) quando $t \rightarrow \infty$.

Observamos agora que se uma solução $(x(t), y(t))$ do Sistema (3.1) deixar as regiões II ou IV, então elas devem cruzar as curvas $y = \phi_1(x)$ ou $y = \phi_2(x)$ e imediatamente entrar na região I ou III. Consequentemente, todas as soluções $(x(t), y(t))$ do Sistema (3.1) que começam na região II ou IV ou nas curvas $y = \phi_1(x)$ ou $y = \phi_2(x)$ devem também tender para a solução de equilíbrio (\bar{x}, \bar{y}) . \square

Para provar o item (b) do teorema, também o retângulo $0 < x < c_1$, $0 < y < c_2$ será dividido em regiões nas quais ambos, \dot{x} e \dot{y} têm sinais fixos. Isto é feito da seguinte maneira. Fazendo $\dot{x} = 0$ no sistema e resolvendo y como função de x temos

$$y = \frac{a_1 x}{b_1(c_1 - x)} \equiv \phi_1(x).$$

Analogamente, fazendo $\dot{y} = 0$ no sistema temos

$$x = \frac{a_2 y}{b_2(c_2 - y)} \text{ ou } y = \frac{b_2 c_2 x}{a_2 + b_2 x} \equiv \phi_2(x).$$

Observe que

- $\phi_1(x)$ e $\phi_2(x)$ são funções monótonas crescentes de x pois

$$\phi_1'(x) = \frac{a_1 b_1 (c_1 - x) + a_1 b_1 x}{(b_1(c_1 - x))^2} = \frac{a_1 b_1 c_1}{(b_1(c_1 - x))^2} > 0$$

e

$$\phi_2'(x) = \frac{b_2 c_2 (a_2 + b_2 x) - b_2^2 c_2 x}{(a_2 + b_2 x)^2} = \frac{a_2 b_2 c_2}{(a_2 + b_2 x)^2} > 0$$

- $\lim_{x \rightarrow c_1} \phi_1(x) = +\infty$ e $\lim_{x \rightarrow +\infty} \phi_2(x) = c_2$
- as curvas $\phi_1(x)$ e $\phi_2(x)$ se interceptam somente no ponto de equilíbrio $(0, 0)$.
- $\phi_1(x)$ cresce mais rapidamente que $\phi_2(x)$ em $x = 0$ pois

$$\phi_2'(0) = \frac{b_2 c_2}{a_2} < \frac{a_1}{b_1 c_1} = \phi_1'(0)$$

já que $a_1 a_2 > b_1 b_2 c_1 c_2$. Logo, $\phi_1(x)$ está acima de $\phi_2(x)$ para $0 < x < c_1$, veja Figura 1.

- $\dot{x}(t) > 0$ para todo (x, y) acima da curva $y = \phi_1(x)$ e negativa em qualquer ponto (x, y) abaixo dessa curva. Analogamente, $\dot{y}(t) > 0$ para todo (x, y) abaixo da curva $y = \phi_2(x)$ e negativa em qualquer ponto (x, y) acima dessa curva. Assim, as curvas $y = \phi_1(x)$ e $y = \phi_2(x)$ dividem o retângulo $0 < x < c_1$, $0 < y < c_2$ em três regiões nas quais \dot{x} e \dot{y} têm sinais fixos.

Também precisaremos dos seguintes lemas para a prova do item (b) do teorema.

Lema 3.7: Toda solução $(x(t), y(t))$ do Sistema (3.1) que começa na região II em $t = t_0$, permanecerá nesta região e tenderá para a solução de equilíbrio $(0, 0)$ quando $t \rightarrow \infty$.

Demonstração: Suponhamos que uma solução $(x(t), y(t))$ do Sistema (3.1), com $(x(t_0), y(t_0))$ na região II, saia dessa região em $t = t^*$. Então $\dot{x}(t^*) = 0$ ou $\dot{y}(t^*) = 0$ já que o único modo da solução deixar a região II é cruzando as curvas $y = \phi_1(x)$ ou $y = \phi_2(x)$. Vamos supor $\dot{x}(t^*) = 0$.

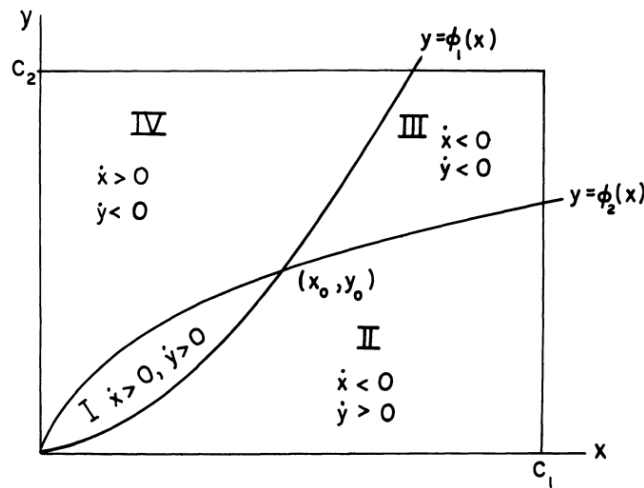


Figura 1 – Item (a) do Teorema.

Derivando ambos os lados da primeira equação do Sistema (3.1) em relação a t e fazendo $t = t^*$ obtemos

$$\ddot{x}(t^*) = b_1(c_1 - x(t^*))\dot{y}(t^*)$$

que é negativo, pois $x(t^*) < c_1$ e $\dot{y}(t^*) < 0$ na curva $\phi_1(x)$, $0 < x < c_1$. Logo t^* é um ponto de máximo de $x(t)$, o que é impossível já que na região II $x(t)$ é sempre decrescente.

Analogamente, se $\dot{y}(t^*) = 0$, então

$$\ddot{y}(t^*) = b_2(c_2 - y(t^*))\dot{x}(t^*)$$

é positivo, pois $y(t^*) < c_2$ e $\dot{x}(t^*) < 0$ na curva $y = \phi_2(x)$, $0 < x < c_1$. Logo t^* é um ponto de máximo de $y(t)$, o que é impossível já que na região II $y(t)$ é sempre decrescente.

Concluimos então que toda solução $(x(t), y(t))$ de (3.1) que começa na região II em $t = t_0$ permanecerá nessa região para $t \geq t_0$. Isso implica que,

- $x(t)$ é monótona decrescente para $t \geq t_0$ com $0 < x(t)$ e
- $y(t)$ é monótona decrescente para $t \geq 0$ com $0 < y(t)$.

Pelo Lema 2.2, $\lim_{t \rightarrow \infty} (x(t), y(t)) = (\xi_1, \xi_2)$ e, pelo Lema 2.3, (ξ_1, ξ_2) é um ponto de equilíbrio do Sistema (3.1). Concluimos que $(\xi_1, \xi_2) = (0, 0)$, o que prova o lema. \square

Lema 3.8: Toda solução $(x(t), y(t))$ do Sistema (3.1) que começa na região I (região III) em $t = t_0$ deve deixá-la num tempo $t = \bar{t}$ (ela deve entrar na região II).

Demonstração: Suponhamos que a solução $(x(t), y(t))$ de (3.1) permaneça na região I (região III) para $t \geq t_0$. Então $x(t)$ é monótona crescente (decrescente) e $y(t)$ é monótona decrescente (crescente) para $t \geq t_0$. Além disso $y(t)$ é menor que c_2 e $x(t)$ é maior que zero, para $t \geq t_0$. Consequentemente, ambas $x(t)$ e $y(t)$ têm limite pelo Lema 2.2, $\lim_{t \rightarrow \infty} x(t) = \xi_1$ e $\lim_{t \rightarrow \infty} y(t) = \xi_2$. Pelo Lema 2.3, o ponto (ξ_1, ξ_2) é de equilíbrio. O único ponto de equilíbrio é o $(0, 0)$. Como

$x(t)$ ($y(t)$) é crescente na região I (região III) concluímos que a solução deve deixar essa região num tempo $t = \bar{t}$. \square

Vamos agora demonstrar o item (b) do Teorema 3.1, veja Figura 2.

Demonstração de 3.1 (b): Pelo Lema 3.8 toda solução que começa nas regiões I ou III deve deixar essa região cruzando as curvas $y = \phi_1(x)$ e $y = \phi_2(x)$, respectivamente. Toda solução que começa nessas curvas deve entrar imediatamente na região II (pois, em $y = \phi_1(x)$ temos $\dot{x} = 0$ e $\dot{y} < 0$, em $y = \phi_2(x)$ temos $\dot{y} = 0$ e $\dot{x} < 0$, o que implica que a solução entra na região II). Pelo Lema 3.7 concluímos que toda solução se aproxima do equilíbrio $(0, 0)$ quando $t \rightarrow \infty$. \square

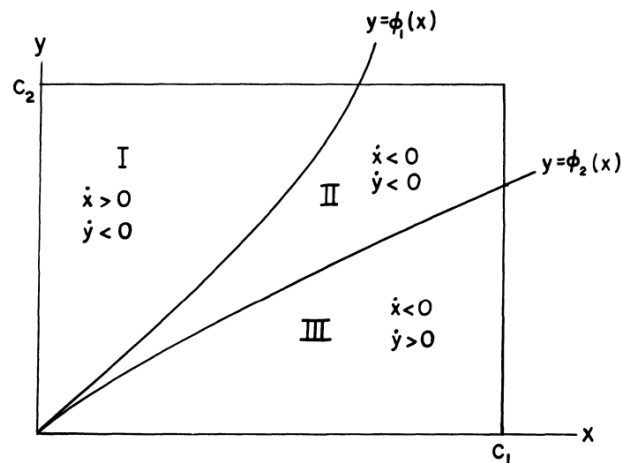


Figura 2 – Item (b) do Teorema.

3.3 Análise local

Nesta seção daremos uma outra prova do Teorema 3.1.

Observemos que o sistema de equações diferenciais (3.1) pode ser escrito na forma matricial $\dot{\mathbf{x}} = A\mathbf{x} + g(\mathbf{x})$, ou seja,

$$\frac{d}{dt} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} -a_1 & b_1 c_1 \\ b_2 c_2 & -a_2 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} - \begin{pmatrix} b_1 x y \\ b_2 x y \end{pmatrix}. \quad (3.2)$$

Para uma prova elegante do item (b) do Teorema 3.1, usando o Teorema de Poincaré-Bendixson, observamos que $\frac{g(\mathbf{x})}{\|\mathbf{x}\|}$ é uma função contínua e que $g(0) = 0$. Podemos então usar o Teorema 2.3 para estudar a estabilidade da solução $(0, 0)$ do Sistema (3.2), ou seja, a estabilidade dessa solução é determinada pela estabilidade da solução $(0, 0)$ do sistema linearizado

$$\frac{d}{dt} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} -a_1 & b_1 c_1 \\ b_2 c_2 & -a_2 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}. \quad (3.3)$$

O polinômio característico da matriz A , dado por $\det(A - \lambda I)$, é

$$\lambda^2 + (a_1 + a_2)\lambda + a_1a_2 - b_1b_2c_1c_2$$

cujas raízes são

$$\lambda = \frac{-(a_1 + a_2) \pm [(a_1 + a_2)^2 - 4(a_1a_2 - b_1b_2c_1c_2)]^{1/2}}{2}.$$

Observe que

$$(a_1 + a_2)^2 - 4(a_1a_2 - b_1b_2c_1c_2) = (a_1 - a_2)^2 + 4b_1b_2c_1c_2 \geq 0$$

e, como $a_1a_2 > b_1b_2c_1c_2$, temos também que

$$(a_1 + a_2)^2 - 4(a_1a_2 - b_1b_2c_1c_2) < (a_1 + a_2)^2.$$

Com isso concluímos que as raízes características são reais e negativas. Logo, a solução de equilíbrio $(0, 0)$ do Sistema (3.2) é assintoticamente estável. Isso implica que qualquer solução $(x(t), y(t))$ do Sistema (3.1) que começa suficientemente próximo da origem $(x, y) = (0, 0)$ tenderá para ela quando $t \rightarrow \infty$.

Agora suponha que uma solução $(x(t), y(t))$ do Sistema (3.1) com $0 < x(t_0) < c_1$ e $0 < y(t_0) < c_2$ não tenda para a origem quando $t \rightarrow \infty$. Pelo que observamos acima, essa solução deve sempre ficar a uma distância mínima da origem. Sendo assim, sua órbita para $t \geq t_0$ pertence à uma região limitada do plano Oxy que não contém pontos de equilíbrio de (3.1). Pelo Teorema de Poincaré-Bendixson, portanto, sua órbita deve espiralar para a órbita de uma solução periódica de (3.1). Mas o sistema (3.1) não tem solução periódica no primeiro quadrante $x \geq 0, y \geq 0$. Isso segue imediatamente do Lema 2.4 (Exercício 11, seção 4.8-Braun) e do fato que

$$\frac{\partial}{\partial x}[-a_1x + b_1(c_1 - x)y] + \frac{\partial}{\partial y}[-a_2y + b_2(c_2 - y)x] = -(a_1 + a_2 + b_1y + b_2x)$$

é estritamente negativo se x e y são não negativos.

Consequentemente, qualquer solução $(x(t), y(t))$ do Sistema (3.1) com $0 < x(t_0) < c_1$ e $0 < y(t_0) < c_2$ se aproxima da solução de equilíbrio $(x, y) = (0, 0)$ quando $t \rightarrow \infty$.

Para a prova da parte (a) do Teorema 3.1, onde $a_1a_2 < b_1b_2c_1c_2$, observamos que o Sistema (3.1) é autônomo bidimensional não-linear da forma

$$\dot{x} = F(x, y) = -a_1x + b_1(c_1 - x)y$$

$$\dot{y} = G(x, y) = -a_2y + b_2(c_2 - y)x$$

que é quase-linear nas vizinhanças dos pontos críticos (\tilde{x}, \tilde{y}) já que as funções $F(x, y)$ e $G(x, y)$ são de classe C^∞ .

Vamos analisar o sistema linear correspondente ao Sistema (3.1) nas vizinhanças dos

pontos críticos (\tilde{x}, \tilde{y}) , ou seja, o sistema

$$\frac{d}{dt} \begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} F_x(\tilde{x}, \tilde{y}) & F_y(\tilde{x}, \tilde{y}) \\ G_x(\tilde{x}, \tilde{y}) & G_y(\tilde{x}, \tilde{y}) \end{pmatrix} \begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} -a_1 - b_1\tilde{y} & b_1c_1 - b_1\tilde{x} \\ b_2c_2 - b_2\tilde{y} & -a_2 - b_2\tilde{x} \end{pmatrix} \begin{pmatrix} u \\ v \end{pmatrix}, \quad (3.4)$$

onde $u = x - \tilde{x}$ e $v = y - \tilde{y}$ e os pontos de equilíbrio a serem analisados são $(0, 0)$ e

$$(\bar{x}, \bar{y}) = \left(\frac{b_1b_2c_1c_2 - a_1a_2}{a_1b_2 + b_1b_2c_2}, \frac{b_1b_2c_1c_2 - a_1a_2}{a_2b_1 + b_1b_2c_1} \right).$$

– $(\tilde{x}, \tilde{y}) = (0, 0)$ é instável.

O Sistema (3.4) em $(0, 0)$ torna-se

$$\begin{pmatrix} \dot{u} \\ \dot{v} \end{pmatrix} = \begin{pmatrix} -a_1 & b_1c_1 \\ b_2c_2 & -a_2 \end{pmatrix} \begin{pmatrix} u \\ v \end{pmatrix},$$

cujos polinômio característico de A é

$$\lambda^2 + (a_1 + a_2)\lambda + a_1a_2 - b_1b_2c_1c_2$$

e suas raízes são

$$\lambda_1 = \frac{-(a_1 + a_2) - [(a_1 + a_2)^2 - 4(a_1a_2 - b_1b_2c_1c_2)]^{1/2}}{2}$$

e

$$\lambda_2 = \frac{-(a_1 + a_2) + [(a_1 + a_2)^2 - 4(a_1a_2 - b_1b_2c_1c_2)]^{1/2}}{2}.$$

Essas raízes são reais com $\lambda_1 < 0$ e $\lambda_2 > 0$ pois

$$a_1a_2 < b_1b_2c_1c_2 \Rightarrow 0 < (a_1 + a_2)^2 - 4(a_1a_2 - b_1b_2c_1c_2) < (a_1 + a_2)^2.$$

– $(\tilde{x}, \tilde{y}) = (\bar{x}, \bar{y})$ é assintoticamente estável.

O Sistema (3.4) em $(0, 0)$ torna-se

$$\begin{pmatrix} \dot{u} \\ \dot{v} \end{pmatrix} = \begin{pmatrix} -a_1 - b_1\bar{y} & b_1c_1 - b_1\bar{x} \\ b_2c_2 - b_2\bar{y} & -a_2 - b_2\bar{x} \end{pmatrix} \begin{pmatrix} u \\ v \end{pmatrix}.$$

Substituindo

$$\bar{x} = \frac{b_1b_2c_1c_2 - a_1a_2}{a_1b_2 + b_1b_2c_2}, \bar{y} = \frac{b_1b_2c_1c_2 - a_1a_2}{a_2b_1 + b_1b_2c_1},$$

temos

$$\begin{pmatrix} \dot{u} \\ \dot{v} \end{pmatrix} = \begin{pmatrix} -b_2c_1 \frac{a_1 + b_1c_2}{a_2 + b_2c_1} & a_1b_1 \frac{b_2c_1 + a_2}{a_1b_2 + b_1b_2c_2} \\ a_2b_2 \frac{b_1c_2 + a_1}{a_2b_1 + b_1b_2c_1} & -b_1c_2 \frac{a_2 + b_2c_1}{a_1 + b_1c_2} \end{pmatrix} \begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} -b_2c_1A & a_1b_1B \\ a_2b_2C & -b_1c_2D \end{pmatrix} \begin{pmatrix} u \\ v \end{pmatrix}$$

onde $A > 0, B > 0, C > 0, D > 0$. As raízes do polinômio característico

$$\lambda^2 + (b_2c_1A + b_1c_2D)\lambda - a_1a_2b_1b_2BC + b_1B - 2c_1c_2AD$$

são

$$\lambda = \frac{-(b_2c_1A + b_1c_2D) \pm [(b_2c_1A + b_1c_2D)^2 - 4(b_1b_2c_1c_2AD - a_1a_2b_1b_2BC)]^{1/2}}{2}.$$

Observe que $AD = 1$ e $BC = \frac{1}{b_1b_2}$ o que implica

$$\lambda = \frac{-(b_2c_1A + b_1c_2D) \pm [(b_2c_1A + b_1c_2D)^2 - 4(b_1b_2c_1c_2 - a_1a_2)]^{1/2}}{2}$$

ou

$$\lambda = \frac{-E \pm \sqrt{E^2 - 4F}}{2}$$

onde $E = b_2c_1A + b_1c_2D$ e $F = b_1b_2c_1c_2 - a_1a_2 > 0$. As raízes são reais pois

$$\begin{aligned} E^2 - 4F &= (b_2c_1A + b_1c_2D)^2 - 4(b_1b_2c_1c_2 - a_1a_2) = \\ &= (b_2c_1A - b_1c_2D)^2 + 4a_1a_2 \geq 0 \end{aligned}$$

e negativas pois $E^2 - 4F < E^2$. Logo, (\bar{x}, \bar{y}) é assintoticamente estável.

Suponhamos que uma solução $(x(t), y(t))$ do Sistema (3.1), com $0 < x(t_0) < c_1$, $0 < y(t_0) < c_2$, não se aproxime do equilíbrio (\bar{x}, \bar{y}) . Pela análise acima, existe uma distância mínima de $(0,0)$ e (\bar{x}, \bar{y}) tal que essa solução deve permanecer. Sendo assim, sua órbita para $t \geq t_0$ pertence à uma região limitada do plano Oxy que não contém pontos de equilíbrio de (3.1). Pelo Teorema de Poincaré-Bendixson, portanto, sua órbita deve espiralar para a órbita de uma solução periódica de (3.1). Mas o sistema (3.1) não tem solução periódica no primeiro quadrante $x \geq 0, y \geq 0$. Isso segue imediatamente do Lema 2.4 (Exercício 11, seção 4.8-Braun) e do fato que

$$\frac{\partial}{\partial x}[-a_1x + b_1(c_1 - x)y] + \frac{\partial}{\partial y}[-a_2y + b_2(c_2 - y)x] = -(a_1 + a_2 + b_1y + b_2x)$$

é estritamente negativo se x e y são não negativos.

Consequentemente, qualquer solução $(x(t), y(t))$ do Sistema (3.1) com $0 < x(t_0) < c_1$ e $0 < y(t_0) < c_2$ se aproxima da solução de equilíbrio $(x, y) = (\bar{x}, \bar{y})$ quando $t \rightarrow \infty$.

3.4 Dados Reais

No caso da gonorreia, é muito difícil avaliar os coeficientes $a_1, a_2, b_1, b_2, c_1, c_2$, é impossível obter até mesmo uma estimativa grosseira de a_2 a qual seria interpretada como o valor médio do tempo em que o sexo feminino permanece infectado (analogamente a_1 seria interpretada como o valor médio do tempo em que o sexo masculino permanece infectado). Isto porque a maioria das mulheres não exibem sintomas. Assim, uma mulher pode ser infecciosa por um tempo variável de um dia para mais de um ano. Entretanto, é ainda possível verificar de dados da saúde pública que $a_1 a_2 < b_1 b_2 c_1 c_2$, como mostramos agora.

Observe que a condição $a_1 a_2 < b_1 b_2 c_1 c_2$ é equivalente a

$$1 < \left(\frac{b_1 c_1}{a_2}\right) \left(\frac{b_2 c_2}{a_1}\right).$$

O valor $\frac{b_1 c_1}{a_2}$ pode ser interpretado como o número médio de homens que uma mulher infectada teve contato durante seu período infeccioso, se todos esses homens eram suscetíveis. Analogamente, o valor $\frac{b_2 c_2}{a_1}$ pode ser interpretado como o número médio de mulheres que um homem infectado teve contato durante seu período infeccioso, se todas essas mulheres eram suscetíveis. Os valores $\frac{b_1 c_1}{a_2}$ e $\frac{b_2 c_2}{a_1}$ são denominados taxas maximais de contato feminino e masculino, respectivamente. O Teorema 3.1 agora pode ser interpretado da seguinte maneira.

- Se o produto das taxas maximais de contato feminino e masculino é maior que um, então a gonorréia se aproximará a um estado estável não nulo.
- Se o produto das taxas maximais de contato feminino e masculino é menor que um, então a gonorréia se aproximará da extinção.

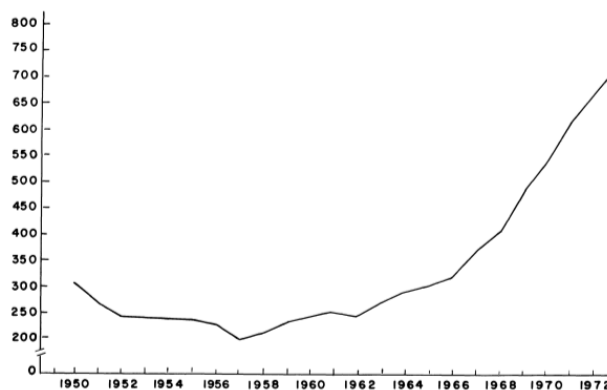


Figura 3 – Número em milhares de casos de gonorreia de 1950 a 1973.

Em 1973, o valor médio de contatos femininos por um homem infectado durante seu período de infecção foi 0,98 enquanto o valor médio de contatos masculinos por uma mulher infectada durante seu período de infecção foi 1,15. Esses valores são aproximações muito boas das taxas

de contato maximais masculina e feminina, respectivamente, e seu produto não excede o produto das taxas maximais de contato masculino e feminino. O número de contatos de um homem ou mulher infectada durante seu período de infecção é levemente menor que as taxas de contato maximais masculinas ou femininas. Além disso, o número real de contatos é frequentemente maior que o número de contatos por um infectado. O produto de 1,15 por 0,98 é 1,127.

Como sugere os dados da Figura 3 poderíamos concluir que a doença gonorreia se propaga rapidamente e de maneira incontrolada, mas pelo o que foi estudado neste capítulo, concluímos que finalmente ela tenderá a um estado estável não nulo.

4. Disseminação da Covid-19

Neste capítulo abordamos a modelagem matemática da disseminação da COVID-19, com ênfase no ajuste para dados coletados do município de São Carlos. Nossa intenção é mostrar como a curva logística pode ajudar a modelar e compreender fenômenos epidemiológicos.

4.1 Como se dá o contágio

Suponhamos que os indivíduos de uma população N tenham contatos entre si. Se X_0 indivíduos estão infectados com um vírus no instante inicial t_0 , então $N - X_0$ é a quantidade de suscetíveis. Queremos deduzir uma equação cuja solução nos permita saber a quantidade de indivíduos infectados em qualquer instante t após o instante t_0 .

Uma hipótese usada em estudos de epidemiologia é que a razão de crescimento em relação ao tempo (velocidade) do número de infectados é proporcional ao número de interações entre infectados e suscetíveis.

Para analisar tais situações, basta colocar uma coluna com os indivíduos infectados e outra com os suscetíveis e, se cada infectado se relaciona (associa) com todos os suscetíveis, deduzimos que o número de interações é o produto X_0 por $N - X_0$.

Se dX/dt é a taxa de crescimento de indivíduos infectados em relação ao tempo t , a hipótese acima se escreve matematicamente

$$\frac{dX}{dt} = K X (N - X),$$

e, percebemos que quanto maior o número de interações (maior $X(N - X)$) mais rápido cresce o número de infectados. A constante de proporcionalidade K representa tanto a possibilidade de contágio em uma interação como a quantidade de interações de um infectado com os suscetíveis. De fato, a relação $\frac{dX}{dt} = K X (N - X)$ já foi explorada em outros contextos de epidemias, conforme discutido por [Rebollo-Perdomo \(2020\)](#) no caso da COVID-19 na Espanha, ilustrando como a taxa de contágio inicial se comporta quase de maneira exponencial.

4.2 Fundamentação Matemática da Curva Logística

O modelo logístico pode ser escrito como a seguinte equação diferencial ordinária (EDO):

$$\frac{dC}{dt} = rC\left(1 - \frac{C}{K}\right), \quad (4.1)$$

em que:

- $C(t)$ representa o número de casos no instante t ;
- $r > 0$ é a taxa intrínseca de crescimento, refletindo a rapidez da disseminação inicial;
- $K > 0$ é o número máximo de casos, de acordo com as limitações inerentes à população.

Solução Analítica

A EDO (4.1) é separável. Partindo de

$$\frac{dC}{dt} = rC\left(1 - \frac{C}{K}\right),$$

podemos rearranjar como

$$\int \frac{dC}{C(K-C)} = \int \frac{r}{K} dt. \quad (4.2)$$

Utilizando a decomposição em frações parciais, escreve-se:

$$\frac{1}{C(K-C)} = \frac{1}{K} \left(\frac{1}{C} + \frac{1}{K-C} \right).$$

Fazendo a integração, obtemos

$$\frac{1}{K} \ln \left| \frac{C}{K-C} \right| = \frac{r}{K} t + \mathcal{C},$$

para algum \mathcal{C} que encapsula a constante de integração. Ao reorganizar termos e tomar exponencial em ambos os lados, encontramos:

$$C(t) = \frac{AK e^{rt}}{1 + A e^{rt}}, \quad (4.3)$$

em que $A = e^{-rt_0}$, resultando em

$$C(t) = \frac{K}{1 + e^{-r(t-t_0)}}. \quad (4.4)$$

O parâmetro t_0 define o **instante de inflexão**, isto é, o tempo em que a taxa de crescimento atinge seu valor máximo. Nesse ponto $t = t_0$, tem-se $C(t_0) = \frac{K}{2}$, indicando que a metade da capacidade máxima K já foi atingida.

A equação logística evidencia três fases típicas de um surto:

- **Crescimento inicial rápido** (quase exponencial), quando $C \ll K$;
- **Ponto de inflexão** ($t = t_0$), onde a taxa de variação \dot{C} é máxima;
- **Saturação**, conforme $t \rightarrow \infty$ e $C(t) \rightarrow K$.

Essa forma sigmoideal de crescimento inspira a análise de modelos epidemiológicos simplificados, como discutimos antes.

4.3 Outro modelo

O número de interações pode diminuir com o tempo, já que vendo a gravidade da doença, muitos indivíduos ficam confinados. Então, para que a equação torne-se mais realística, podemos escrever que a quantidade de suscetíveis é $(1 - C)N - X$, onde C representa a quantidade de confinados da população total (por isso C está entre 0 e 1).

Se C está próximo de 1, então $1 - C$ está próximo de 0, e a consequência disto é que o produto $(1 - C)N$ é pequeno e a quantidade de suscetíveis, $(1 - C)N - X$, que interagem com infectados é pequena. Obtemos então o seguinte modelo

$$\frac{dX}{dt} = K X \left((1 - C)N - X \right).$$

No início dos contágios, não há confinamento e C é próximo de zero, mas à medida que cresce o número de infectados aumenta o número de confinados e C fica próximo a 1. Por exemplo, se as autoridades decretam quarentena no país, então a população gradualmente começa a ficar em casa e somente um pequeno grupo (policiais, médicos, abastecimento básico) continua a interagir. É natural supor que a função C dependendo do tempo tenha um gráfico parecido com a curva logística. Uma das funções mais simples com esta propriedade é

$$C(t) = \frac{C_{\max}}{1 + a e^{bt}},$$

exemplificada por [Rebollo-Perdomo \(2020\)](#), que mostrou como a adoção de medidas de distanciamento social pode ser incorporada como uma função do tempo, restringindo a taxa de contágio em cenários reais.

Por exemplo, suponha que o governo decretou confinamento a partir do vigésimo quinto dia após aumentos consecutivos de infectados. Suponha que no início havia 1% de pessoas confinadas, 50% no dia 25 e foi 90% o confinamento máximo. Então

$$C(t) = \frac{0,9}{1 + a e^{bt}},$$

$C(0) = 1/100$, $C(25) = 50/100$, logo

$$\frac{0,9}{1 + a} = \frac{1}{100}, \quad e \quad \frac{0,9}{1 + a e^{25b}} = \frac{50}{100}.$$

portanto, $a = 89$ e $b = -0,281$. Ou seja,

$$C(t) = \frac{0,9}{1 + 89 e^{-0,281t}}.$$

No caso mais simples em que

$$\frac{dX}{dt} = K X (N - X),$$

que é a equação logística, temos

$$X(t) = \frac{NX_0}{X_0 + (N - X_0)e^{-KNt}}$$

e podemos saber o número de infectados em qualquer instante t , já que podemos resolver a equação e encontrar sua solução explícita. Em contraste com este caso, não conseguimos resolver a equação

$$\frac{dX}{dt} = KX \left((1 - C(t))N - X \right)$$

e obter uma solução em termos de funções elementares. No entanto, Leonhard Euler nos proporcionou a ideia brilhante de resolver numericamente uma equação diferencial. Então com ajuda de um computador, podemos obter o comportamento da solução desta equação e conhecer, aproximadamente, o número de infectados em cada instante t . Abaixo colocamos uma tabela com dados do município de São Carlos do número de infectados pelo vírus da Covid-19 e pudemos perceber, numa aproximação gráfica, novamente a curva logística.

4.4 Aplicação Prática

Metodologia

Os modelos acima nos dão uma boa perspectiva teórica sobre como doenças contagiosas podem se espalhar em uma população. No entanto, para que possamos avaliar sua adequação, é ideal que comparemos os comportamentos esperados com os dados obtidos em casos reais, que sejam suficientemente detalhados e possuam alto grau de confiabilidade.

Visando garantir um bom grau de rigor e a reprodutibilidade de todas as análises, estabelecemos uma metodologia para a obtenção e a curadoria dos dados epidemiológicos. A fonte dos nossos dados será o Painel Coronavírus do Ministério da Saúde ([Ministério da Saúde, 2020](#)), um repositório oficial do governo que disponibiliza informações diárias acerca da evolução da COVID-19 em nível municipal.

A fim de contemplar todo o período disponível e assegurar uma amostra abrangente, decidimos:

- a) Coletar os dados referentes ao histórico da COVID-19, desde os primeiros casos notificados no município até novembro de 2024. O site disponibiliza arquivos .csv compactados com dados de todos os municípios, registrados semanalmente até a última atualização. Escolhemos trabalhar com eles em vez da interface oficial do site pois os dados brutos permitem maior velocidade de processamento e garantem mais confiabilidade nos resultados;
- b) Após obter os dados, verificamos que todos os arquivos continham colunas padronizadas para organizar as informações, como *data*, *estado*, *municipio*, *casosAcumulado*, e *obitosAcumulado*, entre outras. Conferimos se os nomes de campos ou o formato

de datas mantinham-se uniformes e com mesma formatação ao longo dos diferentes arquivos para evitar perda de informações ou problemas de compatibilidade;

- c) Como nosso interesse recai especificamente sobre São Carlos, tanto por limites de processamento quanto para estabelecer o foco onde o trabalho foi desenvolvido, restringimos os dados a entradas cujo campo *municipio* contivesse a string “São Carlos” (ignorando maiúsculas e minúsculas) e cujo campo *regiao* fosse “Sudeste”, prevenindo a inclusão do município homônimo no Estado de Santa Catarina;
- d) Caso a mesma data surgisse repetidamente com valores conflitantes, optamos por preservar o mais atualizado. Se datas apresentassem valores nulos ou claramente fora do esperado (por exemplo, contagens regressivas de casos acumulados), registrávamos em um log de exceções para análise manual.

Buscamos assegurar que todos os dados coletados correspondam fielmente aos registros oficiais, sem lacunas ou repetições. A seguir, abordaremos o desenvolvimento da coleta, filtragem, processamento e análise dos dados, explicando os programas utilizados para cada tarefa. Por fim, criamos uma curva logística com os parâmetros obtidos de nossa análise, e a comparamos com a curva desenhada pelos dados coletados, para verificar a eficácia de nosso modelo.

4.4.1 Coleta e Preparação dos Dados

Após descompactar o arquivo baixado do site do Ministério da Saúde, obtivemos 10 planilhas, contendo dados semanais de 5298 municípios em todo o país. Para separar os dados referentes a São Carlos, usamos o programa abaixo, em Python.

```

1 import pandas as pd
2 import os
3
4 diretorio = r'C:\Users\JMCP\OneDrive\Documentos\TCC-B\
    HIST_PAINEL_COVIDBR_16nov2024 '
5 dados = []
6
7 for arquivo in os.listdir(diretorio):
8     if arquivo.endswith('.csv'):
9         caminho_arquivo = os.path.join(diretorio, arquivo)
10        print(f"Processando arquivo: {arquivo}")
11
12        df = pd.read_csv(caminho_arquivo,
13                        encoding='latin1',
14                        sep=';')
15        dados_sao_carlos = df[
16            (df['municipio'].str.contains('São Carlos', case=False)) &

```

```

17         (df['regiao'] == 'Sudeste')
18     ]
19     print(f"Encontradas {len(dados_sao_carlos)} linhas para São
20           o Carlos neste arquivo")
21     dados.append(dados_sao_carlos)
22
23     dados_combinados = pd.concat(dados, ignore_index=True)
24     caminho_saida = os.path.join(diretorio, 'dados_sao_carlos.csv')
25     dados_combinados.to_csv(caminho_saida, index=False, encoding='
26         latin1', sep=';')
27
28     print(f"Dados combinados salvos em: {caminho_saida}")
29     print(f"Total de linhas encontradas: {len(dados_combinados)}")

```

O processamento dos dados foi realizado seguindo estas etapas:

a) **Importação de bibliotecas:**

```

1 import pandas as pd
2 import os

```

Utilizamos a biblioteca *pandas* para manipulação de dados tabulares, enquanto *os* permite interação com o sistema operacional.

b) **Definição do diretório e estrutura inicial:**

```

4 diretorio = r'C:\Users\JMCP\OneDrive\Documentos\TCC-B\
5           HIST_PAINEL_COVIDBR_16nov2024'
6 dados = []

```

Define o diretório onde estão armazenadas as planilhas com extensão *.csv* (comma-separated values, ou valores separados por vírgula) e cria uma lista vazia para armazenar os dados filtrados.

c) **Iteração pelos arquivos:**

```

7 for arquivo in os.listdir(diretorio):

```

Aqui percorremos todos os arquivos dentro do diretório especificado.

d) **Filtragem das planilhas:**

```

8 if arquivo.endswith('.csv'):

```

Garante que apenas arquivos com extensão *.csv* sejam processados.

e) **Leitura dos dados:**

```

12 df = pd.read_csv(caminho_arquivo,
13                  encoding='latin1',
14                  sep=';')

```

Lê o arquivo CSV e armazena os dados em um DataFrame, utilizando a codificação latin1 e delimitador ;.

f) **Separação dos dados de São Carlos:**

```

15 dados_sao_carlos = df[
16     (df['municipio'].str.contains('São Carlos', case=False)) &
17     (df['regiao'] == 'Sudeste')
18 ]

```

Essa parte filtra os dados que contenham o nome do município "São Carlos" e pertençam à região Sudeste (já que há um município de mesmo nome no Estado de Santa Catarina, na região Sul).

g) **Contagem de registros:**

```

19 print(f"Encontradas {len(dados_sao_carlos)} linhas para São
      Carlos neste arquivo")

```

Exibe a quantidade de registros filtrados para São Carlos.

h) **Armazenamento temporário:**

```

20 dados.append(dados_sao_carlos)

```

Adiciona os dados filtrados à lista previamente criada.

i) **Combinação dos dados:**

```

22 dados_combinados = pd.concat(dados, ignore_index=True)

```

Agrupa os dados de todos os arquivos processados em um único DataFrame. Um DataFrame é uma estrutura de dados bidimensional que pode ser representada matematicamente como uma quádrupla ordenada:

$$DF = (M, L, C, T),$$

onde

- M é uma matriz $m \times n$ de valores
- L é um vetor de índices de linhas de dimensão m
- C é um vetor de rótulos de colunas de dimensão n
- T é um vetor de tipos de dados de dimensão n

j) **Definição do caminho de saída:**

```
23 caminho_saida = os.path.join(diretorio, 'dados_sao_carlos.csv')
```

Define o caminho onde os dados combinados serão salvos.

k) **Salvamento dos dados:**

```
24 dados_combinados.to_csv(caminho_saida, index=False, encoding='
    latin1', sep=';')
```

Salva os dados filtrados em um novo arquivo .csv.

l) **Relatório final:**

```
25 print(f"Dados combinados salvos em: {caminho_saida}")
26 print(f"Total de linhas encontradas: {len(dados_combinados)}")
```

Informa o caminho do arquivo salvo e a quantidade total de registros processados.

4.4.2 Ajuste

Pelo dados obtidos, e o que vimos na literatura, os casos de São Carlos seguiam um comportamento logístico. O ajuste a este modelo foi realizado minimizando a soma dos quadrados dos resíduos (SSR) conforme a Definição 2.4. Para resolver este problema de minimização, utilizamos o método de Levenberg–Marquardt (Definição 2.5), implementado em Python.

Implementação Computacional

O código a seguir implementa o ajuste do modelo logístico aos dados:

```
1 from scipy.optimize import curve_fit
2 from scipy.stats import shapiro
3 import numpy as np
4 import pandas as pd
5
6 def logistica(t, K, r, t0):
7     return K / (1 + np.exp(-r * (t - t0)))
```

```

8 def normalidade(p):
9     return "Distribuição normal" if p > 0.05 else "Distribuição não
        normal"
10
11 file_path = r'C:\Users\JMCP\OneDrive\Documentos\TCC-B\
        HIST_PAINEL_COVIDBR_16nov2024\dados_sao_carlos.csv'
12 df = pd.read_csv(file_path, encoding='utf-8', sep=';')
13
14 dias = pd.to_datetime(df['data']).map(lambda x: (x - pd.to_datetime(df
        ['data'].min())).days)
15 casos = df['casosAcumulado']
16
17 p0 = [max(casos), 0.1, np.median(dias)]
18 params, _ = curve_fit(logistica, dias, casos, p0)
19 K_est, r_est, t0_est = params
20
21 residuos = casos - logistica(dias, K_est, r_est, t0_est)
22 SSR = np.sum(residuos ** 2)
23
24 p_value = shapiro(residuos)[1]
25 res_normalidade = normalidade(p_value)
26
27 def format(label, estimate):
28     return f"{label} estimado: {estimate}"
29
30 resultados = [
31     format("K", K_est),
32     format("r" , r_est),
33     format("t_0" , t0_est),
34     f"Soma dos resíduos quadrados (SSR): {SSR}",
35     f"p-valor: {p_value}",
36     f"Normalidade: {res_normalidade}"
37 ]
38
39 for result in resultados:
40     print(result)

```

Estrutura do Código

```

1  from scipy.optimize import curve_fit
2  from scipy.stats import shapiro
3  import numpy as np
4  import pandas as pd

```

Descrição: Este trecho importa as bibliotecas necessárias para a implementação:

- `curve_fit` do módulo `scipy.optimize` para o ajuste não linear por mínimos quadrados;
- `shapiro` do módulo `scipy.stats` para realizar o teste de Shapiro-Wilk, que verifica se uma amostra segue uma distribuição normal comparando os quantis teóricos com os quantis observados;
- `numpy` para operações numéricas e manipulação de arrays;
- `pandas` para a leitura e manipulação dos dados em formato de planilha (Excel).

```

6  def logistica(t, K, r, t0):
7      return K / (1 + np.exp(-r * (t - t0)))

```

Descrição: Aqui é definida a função logística (4.4)

$$C(t) = \frac{K}{1 + \exp[-r(t - t_0)]},$$

que é a solução analítica da equação diferencial logística

$$\frac{dC}{dt} = rC \left(1 - \frac{C}{K}\right).$$

Os parâmetros são:

- K : o valor máximo de $C(t)$;
- r : a taxa intrínseca de crescimento;
- t_0 : o tempo no qual ocorre o ponto de inflexão (máxima taxa de crescimento).

```

8  def normalidade(p):
9      return "Distribuição normal" if p > 0.05 else "Distribuição não
      o normal"

```

Descrição: Esta função recebe o valor-p resultante de um teste de normalidade e retorna uma mensagem indicando se a hipótese de normalidade dos resíduos pode ser aceita. Se $p > 0.05$, considera-se que os resíduos seguem uma distribuição normal; caso contrário, conclui-se que não seguem.

```

11     file_path = r'C:\Users\JMCP\OneDrive\Documentos\TCC-B\
        HIST_PAINEL_COVIDBR_16nov2024\dados_sao_carlos.csv'
12     df = pd.read_csv(file_path, encoding='utf-8', sep=';')
13
14     dias = pd.to_datetime(df['data']).map(lambda x: (x - pd.
        to_datetime(df['data'].min()))).days
15     casos = df['casosAcumulado']

```

Descrição: Este bloco lê os dados a partir do arquivo CSV dados_são_carlos.csv. Em particular:

- A função `pd.read_csv` carrega os dados para um `DataFrame` denominado `df`, usando codificação UTF-8 e ponto e vírgula como separador;
- A coluna de datas é convertida para o formato `datetime` e mapeada para número de dias desde o primeiro registro;
- A coluna `casosAcumulado` é extraída, correspondendo aos valores observados $C_{\text{obs}}(t_i)$ do número de casos no instante t_i .

```

17     p0 = [max(casos), 0.1, np.median(dias)]
18     params, _ = curve_fit(logistica, dias, casos, p0)
19     K_est, r_est, t0_est = params

```

Descrição: Neste bloco:

- É definido um vetor inicial p_0 usando valores heurísticos baseados nos dados:
 - K inicial é definido como o máximo de casos observados
 - r inicial é definido como 0.1 (valor típico para crescimento epidêmico)
 - t_0 inicial é definido como a mediana dos dias observados
- A função `curve_fit` é usada para ajustar o modelo logístico aos dados $\{(t_i, C_{\text{obs}}(t_i))\}$, ou seja, para encontrar os valores de K , r e t_0 que minimizam a soma dos quadrados dos resíduos, conforme a equação

$$\text{SSR}(K, r, t_0) = \sum_{i=1}^n \left[C_{\text{obs}}(t_i) - \frac{K}{1 + \exp(-r(t_i - t_0))} \right]^2.$$

```

21     residuos = casos - logistica(dias, K_est, r_est, t0_est)
22     SSR = np.sum(residuos ** 2)

```

Descrição: Este bloco calcula e analisa os resíduos do ajuste:

- Os resíduos são calculados como a diferença entre os valores observados (`casos`) e os valores preditos pela função logística ajustada:

$$\varepsilon_i = C_{\text{obs}}(t_i) - \frac{K_{\text{est}}}{1 + \exp[-r_{\text{est}}(t_i - t0_{\text{est}})]}$$

- A soma dos quadrados dos resíduos (SSR) é computada usando `np.sum()`:

$$\text{SSR} = \sum_{i=1}^n \varepsilon_i^2$$

Esta medida quantifica o ajuste global do modelo aos dados, onde valores menores indicam melhor ajuste.

```

24     p_value = shapiro(residuos)[1]
25     res_normalidade = normalidade(p_value)

```

Descrição: Este bloco implementa o teste de Shapiro–Wilk (Definição 2.6) para verificar a normalidade dos resíduos:

- A função `shapiro` do módulo `scipy.stats` é aplicada aos resíduos ε_i , retornando uma estatística de teste e um valor-p;
- O valor-p (segundo elemento do retorno) é extraído e armazenado em `p_value`;
- A função `verificar_normalidade` avalia o valor-p: se maior que 0.05, considera-se que os resíduos seguem distribuição normal; caso contrário, rejeita-se esta hipótese.

Este teste é importante para verificar se os erros do modelo seguem uma distribuição normal, uma das premissas para a validade das estimativas dos parâmetros e seus intervalos de confiança.

```

27     def format(label, estimate):
28         return f"{label} estimado: {estimate}"
29
30     resultados = [
31         format("K", K_est),
32         format("r", r_est),
33         format("t_0", t0_est),
34         f"Soma dos resíduos quadrados (SSR): {SSR}",

```

```

35     f"p-valor: {p_value}",
36     f"Normalidade: {res_normalidade}"
37 ]
38
39 for result in resultados:
40     print(result)

```

Descrição: Este bloco formata e exibe os resultados do ajuste:

- A função `format_result` cria uma string formatada para cada parâmetro e sua incerteza;
- Uma lista `resultados` é criada contendo:
 - Os valores estimados de K , r e t_0
 - A soma dos quadrados dos resíduos (SSR)
 - O p-valor do teste de normalidade
 - A conclusão sobre a normalidade dos resíduos
- Os resultados são impressos sequencialmente usando um laço `for`

Resultados Obtidos:

Após a execução do ajuste do modelo logístico aos dados, os parâmetros estimados foram:

$$K \approx 3.67 \times 10^4, \quad r \approx 7.50 \times 10^{-3}, \quad t_0 \approx 4.49 \times 10^2,$$

$$SSR \approx 9.95 \times 10^9, \quad p\text{-valor} \approx 1.23 \times 10^{-49}.$$

Usando esses parâmetros, vamos comparar graficamente a curva logística obtida com os dados coletados.

4.4.3 Comparando o Modelo com Dados Reais

O código a seguir implementa o processo completo de visualização, gerando a comparação desejada. A seguir, como fizemos acima, iremos descrever o que cada bloco de código faz no programa.

```

1 import numpy as np
2 import matplotlib.pyplot as plt
3 import pandas as pd
4 from datetime import datetime
5 import matplotlib.dates as mdates

```

```

6 try:
7     K = 36722.65596735965
8     r = 0.007502916773061995
9     t0 = 448.69617206306845
10    SSR = 9954864326.08606
11    p_valor = 1.2289754805595289e-49
12
13    superscript = str.maketrans("0123456789-", "0123456789-")
14
15    def format_scientific(num):
16        exp = int(np.floor(np.log10(abs(num))))
17        coef = num / (10**exp)
18        exp_str = str(exp).translate(superscript)
19        return f"{coef:.2f}x10^{exp_str}"
20
21    def logistica(t, K, r, t0):
22        return K / (1 + np.exp(-r * (t - t0)))
23
24    file_path = r"C:\Users\JMCP\OneDrive\Documentos\TCC-B\
25        HIST_PAINEL_COVIDBR_16nov2024\dados_são_carlos.csv"
26    data = pd.read_csv(file_path, sep=';', encoding='latin1')
27
28    data['data'] = pd.to_datetime(data['data'])
29    data['casosAcumulado'] = pd.to_numeric(data['casosAcumulado'],
30        errors='coerce')
31
32    data = data.sort_values('data')
33
34    data_inicio = data['data'].min()
35    data['dias'] = (data['data'] - data_inicio).dt.days
36
37    t_logistica = np.linspace(0, max(data['dias']), 1000)
38    y_logistica = logistica(t_logistica, K, r, t0)
39
40    datas_logistica = data_inicio + pd.to_timedelta(t_logistica, unit='D
41        ')
42
43    plt.figure(figsize=(15, 10))
44    plt.scatter(data['data'], data['casosAcumulado'],
45        color='blue', alpha=0.5, s=30, label='Dados analisados')
46    plt.plot(datas_logistica, y_logistica, 'r-', label='Modelo Logistico
47        ')

```

```

44 def format_data(x, pos=None):
45     date = mdates.num2date(x)
46     meses = {'Jan': 'Jan', 'Feb': 'Fev', 'Mar': 'Mar', 'Apr': 'Abr',
47             'May': 'Mai', 'Jun': 'Jun', 'Jul': 'Jul', 'Aug': 'Ago',
48             'Sep': 'Set', 'Oct': 'Out', 'Nov': 'Nov', 'Dec': 'Dez'}
49     date_str = date.strftime('%b/%Y')
50     for eng, pt in meses.items():
51         date_str = date_str.replace(eng, pt)
52     return date_str
53
54 ax = plt.gca()
55 ax.xaxis.set_major_locator(mdates.MonthLocator(interval=6))
56 ax.xaxis.set_major_formatter(plt.FuncFormatter(format_data))
57
58 plt.xticks(rotation=45)
59
60 plt.xlabel('Data')
61 plt.ylabel('Casos Acumulados')
62 plt.title('Ajuste Logistico para Casos de COVID-19 em Sao Carlos')
63 plt.grid(True, alpha=0.3)
64 plt.legend()
65
66 info_text = (
67     f'K = {format_scientific(K)}\n'
68     f'r = {format_scientific(r)}\n'
69     f't0 = {format_scientific(t0)}\n'
70     f'SSR = {format_scientific(SSR)}\n'
71     f'p-valor = {format_scientific(p_valor)}'
72 )
73
74 plt.text(0.015, 0.95, info_text,
75         transform=plt.gca().transAxes,
76         verticalalignment='top',
77
78         horizontalalignment='left',
79         fontsize=16,
80         bbox=dict(
81             boxstyle='square',
82             facecolor='white',
83             edgecolor='lightgray',
84             alpha=0.9,
85             pad=0.5,

```

```

86         linewidth=0.1
87     ))
88     plt.tight_layout()
89     plt.savefig('ajuste_logistico_temporal.png', dpi=300, bbox_inches='
90         tight')
91     plt.show()
92
93     print("\nInformações sobre os dados:")
94     print(f"Período de análise: {data['data'].min().strftime('%b/%Y')}
95         até {data['data'].max().strftime('%b/%Y')}")
96     print(f"Número total de observações: {len(data)}")
97
98     print("\nResumo dos casos acumulados:")
99     print(data['casosAcumulado'].describe())
100
101 except Exception as e:
102     print(f"\nErro ao executar o código: {str(e)}")
103     import traceback
104     traceback.print_exc()

```

Estrutura do código

Implementação dos Parâmetros O trecho inicial do código define os parâmetros obtidos acima:

```

6     try:
7         # Parametros do modelo logistico
8         K = 36722.65596735965
9         r = 0.007502916773061995
10        t0 = 448.69617206306845
11        SSR = 9954864326.08606
12        p_valor = 1.2289754805595289e-49

```

Estes valores representam o resultado da minimização da soma dos quadrados dos resíduos (SSR).

Processamento Temporal O tratamento dos dados em relação ao tempo envolve várias etapas:

1. Leitura e conversão dos dados:

```

27     data['data'] = pd.to_datetime(data['data'])
28     data['casosAcumulado'] = pd.to_numeric(data['casosAcumulado'],
29         errors='coerce')

```

2. Cálculo do tempo relativo:

```

32 data_inicio = data['data'].min()
33 data['dias'] = (data['data'] - data_inicio).dt.days

```

Este processo estabelece uma origem temporal $t = 0$ correspondente ao primeiro caso registrado.

Visualização dos Dados A construção do gráfico combina dados observados e modelo ajustado:

1. Dados observados são representados por pontos azuis semi-transparentes:

```

41 plt.scatter(data['data'], data['casosAcumulado'],
42 color='blue', alpha=0.5, s=30, label='Dados analisados')

```

2. A curva logística ajustada é plotada como uma linha vermelha contínua:

```

43 plt.plot(datas_logistica, y_logistica, 'r-', label='Modelo Logistico
  ')

```

Localização O código inclui localização para português através de um dicionário de tradução dos meses:

```

44 def format_data(x, pos=None):
45     date = mdates.num2date(x)
46     meses = {'Jan': 'Jan', 'Feb': 'Fev', 'Mar': 'Mar', 'Apr': 'Abr',
47             'May': 'Mai', 'Jun': 'Jun', 'Jul': 'Jul', 'Aug': 'Ago',
48             'Sep': 'Set', 'Oct': 'Out', 'Nov': 'Nov', 'Dec': 'Dez'}
49     date_str = date.strftime('%b/%Y')
50     for eng, pt in meses.items():
51         date_str = date_str.replace(eng, pt)
52     return date_str

```

Por fim, alguns ajustes estéticos foram feitos, e o código gera o gráfico final, salvando-o como uma imagem .png e exibindo-o na tela.

4.4.4 Análise Epidemiológica

A análise dos dados ajustados pela curva logística (vide Figura 4) revela, de modo relativamente claro, três fases distintas no comportamento epidêmico, em linha com os resultados discutidos nesse texto e em abordagens semelhantes:

- a) **Fase Exponencial** ($t \ll t_0$):

Quando t é bem menor que t_0 , o número de casos $C(t)$ pode ser aproximado por

$$C(t) \approx K \exp[r(t - t_0)].$$

No ajuste, isso corresponde aos meses iniciais (aproximadamente até meados de 2021), em que a curva dos dados (pontos azuis na Figura 4) sobe de maneira bastante rápida, embora no gráfico surja alguma dispersão no início. Nessa etapa, os suscetíveis não infectados ainda representam boa parte da população, e a adoção de medidas de distanciamento ainda não é tão intensa.

b) **Fase de Inflexão** ($t \approx t_0$):

O modelo indica que, por volta de $t = 449$ dias, a segunda derivada de $C(t)$ se anula, marcando a *máxima taxa de crescimento* dos casos. Metade do valor assintótico K já teria sido atingido nesse ponto. Na prática, esse momento coincide com um período (cerca de 1 ano e 3 meses após o primeiro registro) em que o ritmo de novos casos começa a arrefecer. Não é de todo absurdo supor que a intensificação das medidas de controle (uso de máscaras, testagens mais frequentes) aumento de pessoas isoladas e menor fluxo populacional contribuíram para essa redução na taxa de crescimento.

c) **Fase de Saturação** ($t \gg t_0$):

Em tempos muito posteriores a t_0 , a curva tende ao patamar

$$\lim_{t \rightarrow \infty} C(t) = K.$$

Observa-se que o modelo previu um valor em torno de 3.67×10^4 de casos acumulados. Comparando com os dados mais recentes (até meados de 2024), o número de casos reais apresenta flutuações, mas aproxima-se dessa faixa, ainda que haja pequenas discrepâncias. Esse regime de saturação ocorre quando o grupo de suscetíveis e a taxa de contato efetiva não favorecem mais o crescimento explosivo.

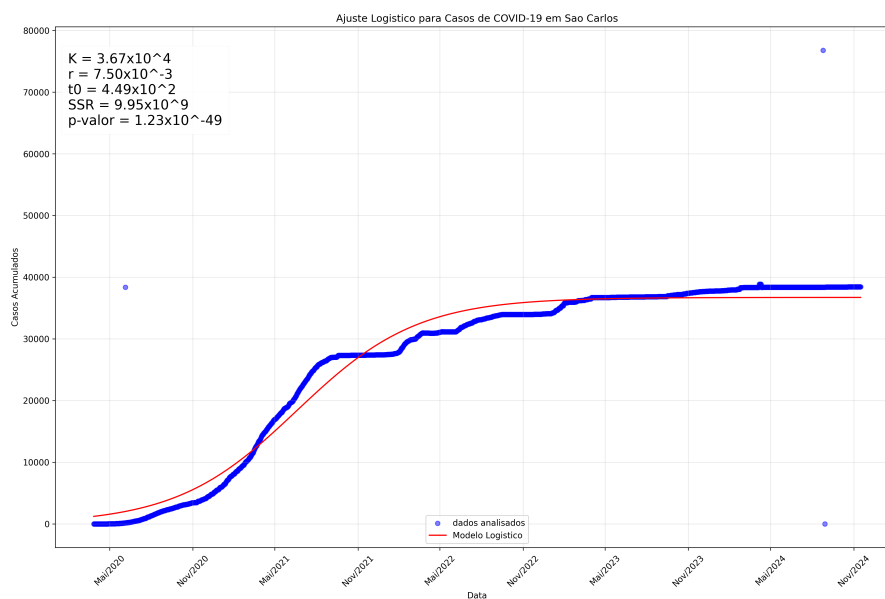


Figura 4 – Ajuste Logístico para Casos de COVID-19 em São Carlos

Essas fases são diretamente influenciadas por *medidas de controle sanitário* (distanciamento social, testagem, vacinação) e pela *adesão da população*, conforme apontado em Nicola (2020). Um leve decréscimo do parâmetro r (taxa de crescimento) ou uma redução efetiva no valor de K (via imunização em larga escala) podem atrasar a chegada ao pico ou diminuir o ápice de casos.

O valor $r \approx 0.0075$ implica uma velocidade de crescimento moderada no começo, distinta de cenários muito agressivos em que r poderia chegar a 0.02 ou 0.03. Pequenas variações nesse parâmetro mudariam significativamente o momento do pico.

Embora a SSR de 9.95×10^9 seja elevado, esta é relativizada pela ordem de grandeza dos dados (dezenas de milhares de casos), de modo que, na escala total, o ajuste é razoavelmente bom.

O p-valor do teste de Shapiro–Wilk ($\approx 1.23 \times 10^{-49}$) nos indica a falta de normalidade dos resíduos, evidenciando que há flutuações diárias e efeitos sazonais que o modelo logístico puro não capta. Como vimos no período de isolamento social da pandemia, isso claramente é verdade, com novas variantes como Delta e Ômicron surgindo e diferentes políticas públicas sendo estabelecidas, que sugeriam diferentes medidas sanitárias, métodos preventivos e tratamentos, trazendo grande variabilidade aos dados.

5. Comentários

Nosso modelo para a gonorreia é um pouco grosseiro já que consideramos todos os homens e mulheres promíscuos independentemente da idade. Um modelo mais preciso pode ser obtido separando por idade as populações masculinas e femininas em diferentes grupos e, então, calcular a taxa de crescimento de infecciosos em cada um desses grupos. Tem estudos sobre este caso e sua análise é mais difícil do que a que fizemos aqui, mas um resultado completamente análogo ao Teorema 2.3 foi obtido: ou a gonorreia tende à extinção em cada grupo ou se aproxima de uma constante positiva. No capítulo 4, podemos dizer que a forma da curva C depende das medidas do governo e do comportamento da população. Por exemplo, depende do dia em que o governo decreta quarentena e da quantidade de pessoas que ficam confinadas. Com a experiência que tivemos com a pandemia da Covid-19 e com modelos matemáticos estudados, podemos afirmar que a inexistência de vacina e considerando as dificuldades naturais de implementar rapidamente um tratamento clínico, resta uma única estratégia eficaz de combater uma epidemia, diminuir as interações sociais pelo distanciamento social e a difusão de proteção pessoal, como uso de álcool gel e máscaras. Para modelar a disseminação da gonorreia, no caso de relações homossexuais, caso este que poderia ser adaptado para a Covid-19, chegamos à equação

$$\dot{x}(t) = -\alpha_1 x + \beta_1 x(c - x),$$

que é a equação logística

$$\dot{x}(t) = (-\alpha_1 + \beta_1 c)x - \beta_1 x^2,$$

onde c é a quantidade total de indivíduos homossexuais, $x(t)$ o número de infectados nessa população, α_1 a taxa com que são curados e β_1 a taxa de novos infectados. Pelo que estudamos no TCC-A, podemos concluir que, se $\alpha_1 < \beta_1 c$, então

$$x(t) \rightarrow \frac{\beta_1 c - \alpha_1}{\beta_1}, \quad \text{quando } t \rightarrow \infty,$$

ou seja, o número de infectados tende a um valor fixo não nulo.

A curva logística obtida em nossa análise oferece uma boa visão das fases da epidemia (com crescimento, pico e estabilização bem definidos), mas não é a única forma de modelar a propagação da Covid-19, nem mesmo a melhor, pois não incorpora fatores como re-infecção, subnotificação ou heterogeneidade geográfica, que são de grande importância ao se tratar de uma pandemia. Outros modelos podem ser utilizados, como os modelos SIR/SEIR, especialmente ao mirarmos em projeções de longo prazo ou cenários com variação abrupta nas medidas sanitárias. Ainda assim, a curva logística ajustada nos ajuda a compreender em que estágio do surto a localidade se encontra, auxiliando a direcionar decisões de saúde pública.

Referências

BRAUN, Martin. **Differential Equations and Their Applications**. 4th. [S. l.]: Springer, 1993. Citado nas pp. 10, 18.

MINISTÉRIO DA SAÚDE. **COVID-19: Portal do Ministério da Saúde**. [S. l.: s. n.], 2020. Disponível em: <https://covid.saude.gov.br/>. Acesso em: 16 nov. 2024. Citado na p. 39.

NICOLA, Selma. A matemática e a epidemia. **Professor de Matemática Online**, v. 8, p. 298–306, 2020. Citado na p. 54.

NOCEDAL, Jorge; WRIGHT, Stephen J. **Numerical Optimization**. 2nd. [S. l.]: Springer, 2006. Citado na p. 10.

REBOLLO-PERDOMO, Salomón. Un modelo simple para el número de infectados por Covid-19. **Materials matemàtics**, p. 1–12, 2020. Disponível em: <https://ddd.uab.cat/record/222029>. Citado nas pp. 36, 38.

SHAPIRO, Samuel S.; WILK, Martin B. An Analysis of Variance Test for Normality (Complete Samples). **Biometrika**, v. 52, p. 591–611, 1965. Citado na p. 10.

Exceto quando indicado o contrário, a licença deste item é descrito como
Attribution-NonCommercial-NoDerivs 3.0 Brazil

