

*Id :*  
*deras.cls, v1.272004/04/0114 :*  
*07 :*  
*46mvsfrassonExp*

# **Livros Grátis**

<http://www.livrosgratis.com.br>

Milhares de livros grátis para download.

538 Levartoski de Araújo, Aderaldo Irineu  
S477t Aninhamento em Redes Bipartidas.  
Aderaldo Irineu Levartoski de Araújo. Natal: 2009.  
Tese de doutorado - Universidade Federal do Rio Grande do Norte.  
Orientador: Gilberto Corso.  
Área de Concentração: Física da Matéria Condensada.  
1. Física - Tese. 2. Redes Complexas. 3. Padrões em Redes.  
4. Redes Bipartidas 5. Interações interespecíficas. 6. Análise de Metacomunidades.  
I. Corso, Gilberto. II. Universidade Federal do Rio Grande do Norte. Departamento de Física Teórica e Experimental.  
III. Título.

UNIVERSIDADE FEDERAL DO RIO GRANDE DO NORTE  
CENTRO DE CIÊNCIAS EXATAS E DA TERRA  
DEPARTAMENTO DE FÍSICA TEÓRICA E EXPERIMENTAL

CURSO DE PÓS - GRADUAÇÃO EM FÍSICA

# *Aninhamento em Redes Bipartidas*

Aderaldo Irineu Levartoski de Araújo

Tese de Doutorado

Orientador:  
Prof. Dr. Gilberto Corso

Natal  
Dezembro de 2009

UNIVERSIDADE FEDERAL DO RIO GRANDE DO NORTE  
CENTRO DE CIÊNCIAS EXATAS E DA TERRA  
DEPARTAMENTO DE FÍSICA TEÓRICA E EXPERIMENTAL  
CURSO DE PÓS - GRADUAÇÃO EM FÍSICA

**Aderaldo Irineu Levartoski de Araújo**

## *Aninhamento em Redes Bipartidas*

Tese apresentada ao Curso de Pós-Graduação em Física da Universidade Federal do Rio Grande do Norte como parte dos requisitos para a obtenção do título de Doutor em Física.

---

Prof. Dr. Gilberto Corso  
Orientador, UFRN

---

Prof. Dr. Luciano Rodrigues da Silva  
UFRN

---

Prof. Dr. Adriana Monteiro de Almeida  
UFRN

---

Prof. Dr. José Garcia Vivas Miranda  
UFBA

---

Prof. Dr. Paulo Henrique Ribeiro Barbosa  
UFPI

Aprovada em Dezembro de 2009.

*A meus filhos Bianca e Nikolas*

# *Agradecimentos*

À minha família.

Ao Prof. Gilberto Corso pelo trabalho de orientação, por sua paciência e amizade.

Ao Prof. José Luiz de Attayde, à Prof. Adriana Monteiro de Almeida e ao Prof. Márcio Zicán Cardoso por me ajudarem com conceitos de Ecologia. Aos professores, colegas e funcionários do Departamento de Física Teórica e Experimental e do Departamento de Biofísica da UFRN por contribuírem com meu aprimoramento pessoal e profissional.

”O presente trabalho foi realizado com o apoio do Conselho Nacional de Desenvolvimento Científico e Tecnológico - CNPq - Brasil”.

# *Resumo*

Apresentamos um índice de aninhamento que mede o padrão de aninhamento de redes bipartidas, um problema que surge em ecologia teórica. Nossa medida é construída através da soma das distâncias dos elementos ocupados na matriz de adjacência da rede. Este índice quantifica diretamente o desvio de uma dada matriz em relação a um padrão aninhado. No caso mais simples a distância do elemento  $a_{i,j}$  da matriz é  $d_{i,j} = i + j$ , a distância de Manhattan. Uma distância genérica é obtida através de  $d_{i,j} = (i^\chi + j^\chi)^{1/\chi}$ . O índice de aninhamento é definido por  $\nu = 1 - \tau$ , onde  $\tau$  é a "temperatura" da matriz. Construímos o índice de temperatura utilizando dois padrões de referência: a distância da matriz completamente aninhada, que corresponde à temperatura zero, e a distância da matriz aleatória média, definida de modo que sua temperatura seja um. Discutimos uma importante característica do problema, a ocupação  $\rho$  da matriz. Abordamos esta questão introduzindo o índice métrico  $\chi$  que permite o ajuste de matrizes com diferentes ocupações.

# *Abstract*

We present a nestedness index that measures the nestedness pattern of bipartite networks, a problem that arises in theoretical ecology. Our measure is derived using the sum of distances of the occupied elements in the adjacency matrix of the network. This index quantifies directly the deviation of a given matrix from the nested pattern. In the most simple case the distance of the matrix element  $a_{i,j}$  is  $d_{i,j} = i + j$ , the Manhattan distance. A generic distance is obtained as  $d_{i,j} = (i^\chi + j^\chi)^{1/\chi}$ . The nestedness index is defined by  $\nu = 1 - \tau$  where  $\tau$  is the "temperature" of the matrix. We construct the temperature index using two benchmarks: the distance of the complete nested matrix that corresponds to zero temperature and the distance of the average random matrix that is defined as temperature one. We discuss an important feature of the problem: matrix occupancy  $\rho$ . We address this question using a metric index  $\chi$  that adjusts for matrix occupancy.

# Sumário

## Lista de Figuras

<b>Introdução</b>	<b>1</b>
<b>1 Motivação Ecológica</b>	<b>4</b>
1.1 Redes Bipartidas . . . . .	4
1.2 Teias Bipartidas de Interações Mutualísticas . . . . .	8
1.3 Teias Bipartidas de Interações Herbívoro-Hospedeira . . . . .	8
1.4 Redes de Mutualismo e Herbivoria . . . . .	9
1.5 Teias Bipartidas em Biogeografia e Ecologia de Conservação . . . . .	10
<b>2 Conceitos Matemáticos</b>	<b>13</b>
2.1 Grafos . . . . .	13
2.1.1 Grafos orientados como modelos matemáticos . . . . .	15
2.1.2 Redes como modelos matemáticos . . . . .	16
2.1.3 O grau de um vértice . . . . .	18
<i>Teorema 2.1</i> . . . . .	19
<i>Teorema 2.2</i> . . . . .	19
2.1.4 Grafos isomorfos . . . . .	19
<i>Teorema 2.3</i> . . . . .	21
<i>Teorema 2.4</i> . . . . .	22
2.1.5 Grafos Conectados . . . . .	22
2.1.6 Vértices Cortados e Pontes . . . . .	24
<i>Teorema 2.5</i> . . . . .	25
2.2 O problema da ponte de Königsberg . . . . .	25
<i>Teorema 3.1</i> . . . . .	26
2.3 O problema da dança: Grafos Bipartidos . . . . .	26
<i>Teorema 4.1</i> . . . . .	28
2.4 Grafos e Matrizes . . . . .	28
2.5 Conceitos Métricos . . . . .	30
2.6 Geometria de Manhattan ou <i>taxicab</i> . . . . .	32

<b>3</b>	<b>Desenvolvimento Analítico do Índice de Aninhamento</b>	<b>33</b>
3.1	Aninhamento em Matrizes de Adjacência . . . . .	33
3.2	Temperatura Manhattan . . . . .	34
3.3	Ocupação da Matriz e Otimização da Métrica . . . . .	37
3.4	Aplicação do novo índice de aninhamento a dados de campo . . . . .	41
	<b>Considerações Finais</b>	<b>43</b>
	<b>Apêndice A - Código do Programa MetricNest</b>	<b>45</b>
	<b>Referências</b>	<b>53</b>

# *Lista de Figuras*

1	Um exemplo de dois padrões de referência usados para encontrar o índice de aninhamento: a matriz perfeitamente aninhada à esquerda e a matriz aleatória à direita. Em ambos os casos temos $N = 12$ sítios ocupados e $L_1 = L_2 = 5$ . . . . .	3
1.1	Grafo bipartido representando uma rede de herbivoria com 13 espécies de plantas (A a M) e 13 espécies de insetos endófagos (1 a 13). Cada vez que um inseto emerge de uma flor uma conexão entre ambos é estabelecida e representada por uma linha. Em outras palavras, este inseto particular utiliza esta planta particular como local de alimentação. As informações são sintetizadas pela da matriz de adjacência. Dados coletados em uma região montanhosa do Sudeste do Brasil [1]. . . . .	6
1.2	Estruturas mais comuns em matrizes. Matriz aninhada à esquerda. Note que os insetos mais generalistas (primeiras colunas) interagem com as plantas mais generalistas (primeiras linhas). Matriz compartimentada no centro. Espécies dentro de um compartimento interagem mais fortemente entre si que com de outros compartimentos. Matriz aleatória à direita. Não há uma estrutura bem definida e as espécies interagem ao acaso. . . . .	7
2.1	Diagramas representando grafos . . . . .	14
2.2	Grafo orientado ou dígrafo . . . . .	15
2.3	Dígrafos simétricos . . . . .	16
2.4	Grafos representando redes . . . . .	16
2.5	Grafos assinalados . . . . .	17
2.6	Grafos com valores funcionais inteiros . . . . .	17
2.7	Loop-dígrafo . . . . .	18
2.8	Vértices com diversos graus . . . . .	19
2.9	Grafos completos . . . . .	20
2.10	Grafos isomorfos . . . . .	20
2.11	Grafos não isomorfos de mesma ordem e tamanho . . . . .	21
2.12	Classes de isomorfismo . . . . .	22
2.13	Subgrafo $H$ conectado ao grafo $G$ . . . . .	22

2.14	Grafo com quatro componentes . . . . .	23
2.15	Ciclos e circuitos . . . . .	23
2.16	Vértices cortados e pontes . . . . .	24
2.17	Multigrafo representando as pontes de Königsberg . . . . .	25
2.18	Grafos bipartidos . . . . .	27
2.19	Vértices emparelhados . . . . .	27
2.20	Grafo $G$ e sua matriz de adjacência . . . . .	29
2.21	Grafo da figura anterior com rotulação diferente . . . . .	29
2.22	Matriz de incidência . . . . .	30
2.23	Distâncias de Minkowski para $\chi = 1$ (linhas sólidas) ou $\chi = 2$ (linhas tracejadas). . . . .	32
3.1	A representação de um padrão perfeitamente aninhado com área $\rho$ no quadrado unitário. A linha oblíqua delimita o padrão aninhado. A equação Cartesiana desta linha é indicada na figura. . . . .	36
3.2	O comportamento de $d_{nest}$ e $d_{rand}$ versus $\rho$ . As curvas obedecem à regra $d_{nest} < d_{rand}$ , como esperado. Por construção, essas curvas não dependem do tamanho da matriz. . . . .	38
3.3	Estas duas matrizes estão igualmente bem aninhadas. Escolhemos um parâmetro métrico $\chi$ adequado para produzir índices "similares" de aninhamento. . . . .	38
3.4	As curvas $1 = (x^\chi + y^\chi)^{1/\chi}$ para vários valores de $\chi$ indicados na figura. Escolhemos o melhor $\chi$ de modo que delimite o padrão completamente aninhado para uma dada ocupação. Esta figura deve ser comparada com a figura 3.3. Nossa estratégia é encontrar um $\chi$ ótimo de modo que a curva delimite o padrão aninhado. . . . .	39
3.5	O comportamento de $d_{nest}$ e $d_{rand}$ versus $\rho$ . As curvas obedecem a regra $d_{nest} < d_{rand}$ como esperado. Alguns valores de $\chi$ correspondentes a $\rho$ estão indicados no topo da figura. . . . .	40

# *Introdução*

Um dos principais objetivos da ecologia de comunidades e da biogeografia é explicar o padrão não aleatório da distribuição de espécies entre sítios, que podem ser ilhas, fragmentos de vegetação ou uma mancha de recursos. Na procura por estes padrões alguns estudos têm descrito uma estrutura aninhada: espécies em sítios pobres são um subconjunto de espécies em sítios mais ricos, em função de migração ou extinção seletivas [2]. O conceito de aninhamento foi primeiramente proposto para explicar a composição de espécies em ilhas, onde ilhas pobres em espécies são um subconjunto de ilhas mais ricas. Mais recentemente, o conceito de aninhamento tem sido também usado para descrever interações interespecíficas entre comunidades, como entre plantas e animais [3]. Padrões aninhados em matrizes de interação entre espécies emergem quando espécies especialistas interagem com espécies que formam subconjuntos bem definidos das espécies com as quais os generalistas interagem [4]. Em ambos os contextos o aninhamento é um padrão, ou um índice, que é estimado usando a matriz de presença e ausência. Embora matrizes de interação e matrizes biogeográficas sejam similares em muitos aspectos, elas diferem na dinâmica subjacente aos processos e nos padrões resultantes [3, 5].

Muitos estudos nesta área até o presente prescindem de uma formulação matemática do conceito de aninhamento. Construções analíticas são raras e há poucas demonstrações em artigos que tratam de aninhamento. De modo diferente, seguindo uma forte tradição nesta área da ecologia, encontramos um grande uso de simulações Monte Carlo e de procedimentos de reamostragem, baseados em avaliações estatísticas e testes de hipótese. O recente interesse no assunto reflete-se na proliferação de diferentes estimadores de aninhamento. Todos eles são manipulações da matriz de presença e ausência e vêm sendo testados contra uma variedade de modelos nulos [6, 7, 8]. Dois artigos muito interessantes, que vão além da maioria dos trabalhos na área e apresentam o aninhamento em uma perspectiva dinâmica são [9, 10].

A descrição em termos de redes do problema da biogeografia de ilhas, uma área da ecologia de metacomunidades, é a seguinte. Dado um conjunto  $I_i$  de ilhas e um conjunto  $S_j$  de espécies, cada vez que a espécie  $i$  é encontrada na ilha  $j$  uma ligação entre ambos os conjuntos é estabelecida. Por outro lado, no contexto da ecologia de comunidades, a rede bipartida é uma rede de interação. Em termos gerais, encontramos dois tipos de interações:

antagonistas e mutualistas. O exemplo típico de rede de antagonistas é a teia alimentar, onde temos herbívoros e carnívoros em um dado ecossistema. O padrão clássico de rede mutualista é formado por plantas com flores e seus pássaros ou insetos polinizadores. Em todos estes casos o aninhamento é uma ferramenta essencial para caracterizar a estrutura das redes.

Em nosso trabalho propomos que o aninhamento possa ser especificado analiticamente e que o índice de aninhamento possa ser bem definido não apenas de forma algorítmica, mas possa também ser matematicamente bem definido. Apresentamos um novo índice para quantificar o aninhamento que é baseado em propriedades métricas da matriz de adjacência. Para estabelecer o problema de modo rigoroso iniciamos com dois pressupostos. Primeiro, um padrão perfeitamente aninhado é constituído por uma estrutura aninhada de sítios pobres em espécies inseridos em sítios mais ricos, ou de grupos de espécies de distribuição geográfica restrita dentro de grupos de espécies de distribuição mais ampla. Segundo, um índice efetivo de aninhamento precisa quantificar o quanto uma dada matriz desvia-se do padrão perfeitamente aninhado.

Baseados nestes dois pressupostos e com a ajuda de padrões de referência bem definidos nós definimos uma métrica que chamamos temperatura  $\tau$  e um índice de aninhamento  $\nu = 1 - \tau$ . Seguimos a analogia com a temperatura feita por Atmar and Patterson [11], mas nosso índice é analiticamente mais próximo ao índice  $N_0$  de Atmar e Patterson [12] e ao índice de discrepância de Brualdi e Sanderson [13], como veremos no capítulo 3.

A noção de temperatura inspirou o desenvolvimento de nosso índice de aninhamento. Quando um sistema termodinâmico está a temperatura zero todas as partículas estão no estado de mínima energia, ou no estado fundamental. É natural associar ao estado completamente aninhado a temperatura zero. Extendendo a analogia com a termodinâmica, para definir a métrica temperatura faz-se necessário um segundo ponto de referência bem definido. Para este ponto usamos a média de um conjunto de matrizes randômicas uniformemente distribuídas, cuja temperatura é definida como sendo um. Estes dois padrões de referência são visualizados na figura 1.

Nosso trabalho ocupa-se de uma nova medida de aninhamento que é baseada em propriedades métricas da matriz e é organizado como segue. No capítulo 2 apresentamos os aspectos da teoria de grafos e da teoria de redes que serão utilizados ou estão intimamente relacionados com o trabalho desenvolvido. No capítulo 3 fazemos uma descrição detalhada do problema e o colocamos no contexto de estudos prévios. No capítulo 4 desenvolvemos os índices de temperatura e de aninhamento usando a distância de Manhattan. No capítulo 5 ampliamos o desenvolvimento teórico para incluir a comparação

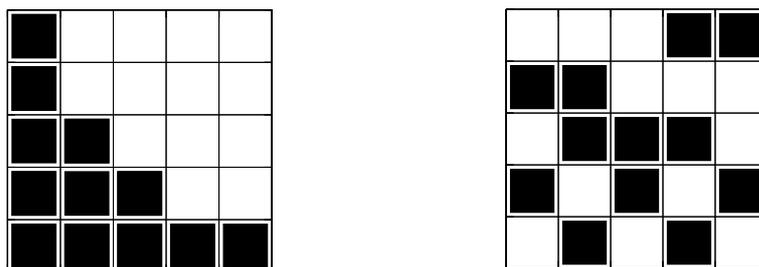


Figura 1: Um exemplo de dois padrões de referência usados para encontrar o índice de aninhamento: a matriz perfeitamente aninhada à esquerda e a matriz aleatória à direita. Em ambos os casos temos  $N = 12$  sítios ocupados e  $L_1 = L_2 = 5$ .

entre padrões de aninhamento com diferentes níveis de ocupação. Na seção 6 oferecemos nossas considerações finais.

# *1 Motivação Ecológica*

## **1.1 Redes Bipartidas**

Teias tróficas são constituídas por populações que se alimentam uns dos outros, e seu estudo é de fundamental importância para o desenvolvimento de teorias ecológicas, aprimorando a habilidade de investigação e previsão de processos fundamentais para a estrutura de todos os sistemas ecológicos, dos quais fazem parte os consumidores e seus recursos alimentares [14]. Desta forma, a busca de padrões gerais em teias tróficas pode levar ao desenvolvimento de teorias gerais, capazes de prever relações entre aspectos estruturais, como riqueza de espécies, e funcionais, como interações tróficas em comunidades [15]. Poucas teias tróficas existentes apresentam interações descritas para espécies biológicas (ou morfoespécies ou OTUs - unidades taxonômicas operacionais). Ao invés de espécies, é comum o uso de "espécies tróficas" ou "trofoespécies", grupos funcionais formados por espécies, não necessariamente aparentadas, que dividem as mesmas presas e predadores [16]. O uso de trofoespécies pode parecer mais prático do ponto de vista metodológico, por não haver necessidade de identificação de todas as espécies e desta forma tanto reduzir tendências causadas por erros na identificação de alguns grupos, como tornar a teia mais tratável, diminuindo sua complexidade [17, 15]. Por outro lado, na base de teias tróficas trofoespécies podem conter centenas de espécies, como por exemplo a trofoespécie de microorganismos, enquanto vertebrados geralmente são descritos como espécies biológicas e grupos como invertebrados apresentam resolução taxonômica intermediária. O agrupamento de espécies em trofoespécies é subjetivo e depende do autor do estudo e de seus objetivos, o que pode comprometer em muito a comparação entre teias. Desta forma, o uso de trofoespécies é controverso e pode alterar alguns parâmetros das teias [18]. As teias tróficas resolvidas no nível específico são poucas e somente recentemente se deu maior atenção à sua estrutura. Martinez [19] mostrou que atributos de teias resolvidas em nível de espécie são mais consistentes para cada teia individual, embora padrões gerais sejam mais frouxos, indicando que teias com trofoespécies apresentam estrutura mais simples, das quais padrões gerais são obtidos com maior facilidade. Entretanto, como estes padrões

gerais deixam de existir em teias melhor definidas, a conclusão é que estes padrões obtidos em teias de trofoespécies podem não refletir a realidade e sim ser consequência de sua pouca resolução. Com exceção de comunidades extremamente simples, informações detalhadas sobre todas as espécies e suas interações tróficas são praticamente impossíveis de serem obtidas. Entretanto, estudos detalhados de guildas de consumidores ou de taxocenoses em uma comunidade são obtidos com maior facilidade [20]. Nos últimos anos, tem sido mais freqüente o estudo de teias com apenas dois ou três níveis tróficos, mas onde todas as espécies estão identificadas e algumas vezes suas interações quantificadas [21, 15, 22, 23, 24, 25]. Em uma teia bipartida somente dois níveis tróficos são analisados, e sua representação gráfica pode dar-se tanto na forma de uma teia quanto de uma matriz de associação, onde todas as espécies podem ser conhecidas e suas interações estudadas em maior detalhe. Uma rede bipartida qualitativa de interações herbívoro-hospedeira pode ser representada também por uma matriz com as espécies de plantas nas linhas e as espécies de insetos nas colunas. As células na matriz são preenchidas com "1" se houver interação e "0" onde não houver interação. A figura 1.1 mostra uma rede deste tipo e a matriz de adjacência que a representa.

Matrizes de interação e sua representação gráfica na forma de teias bipartidas podem estar organizadas na forma de aninhamento ou compartimentação, ou simplesmente não estar organizadas, apresentando distribuição aleatória de interações. Matrizes aninhadas apresentam um tipo de distribuição onde cada espécie interage apenas com um subconjunto de espécies mais generalistas (figura 1.2 à esquerda). Desta forma, as espécies mais generalistas tanto de animais quanto de hospedeiras interagem entre si formando um "core" denso de interações, ao qual todo o restante da comunidade está ligado [26, 27]. No caso de ecologia de conservação, quando há a fragmentação de uma mata, por exemplo, várias espécies são extintas localmente por não conseguir viver em um ambiente em pedaços. Estas extinções não são aleatórias uma vez que dependem das especificidades do ecossistema. Matrizes aninhadas são pouco sensíveis à retirada aleatória de espécies, mas muito sensíveis à retirada de espécies muito conectadas. São também menos susceptíveis à perda de espécies por fragmentação [28, 29].

Matrizes compartimentadas apresentam-se divididas em subseções, ou compartimentos, de forma que espécies em um compartimento interagem mais fortemente entre si que com espécies de outros compartimentos (figura 1.2 no centro) [30, 31]. Uma estrutura compartimentada aumenta a estabilidade do sistema pois os compartimentos isolam efeitos em cascata após a exclusão aleatória de espécies [32, 29]. Matrizes de interações também podem apresentar um padrão aleatório, se cada espécie de planta (ou animal)

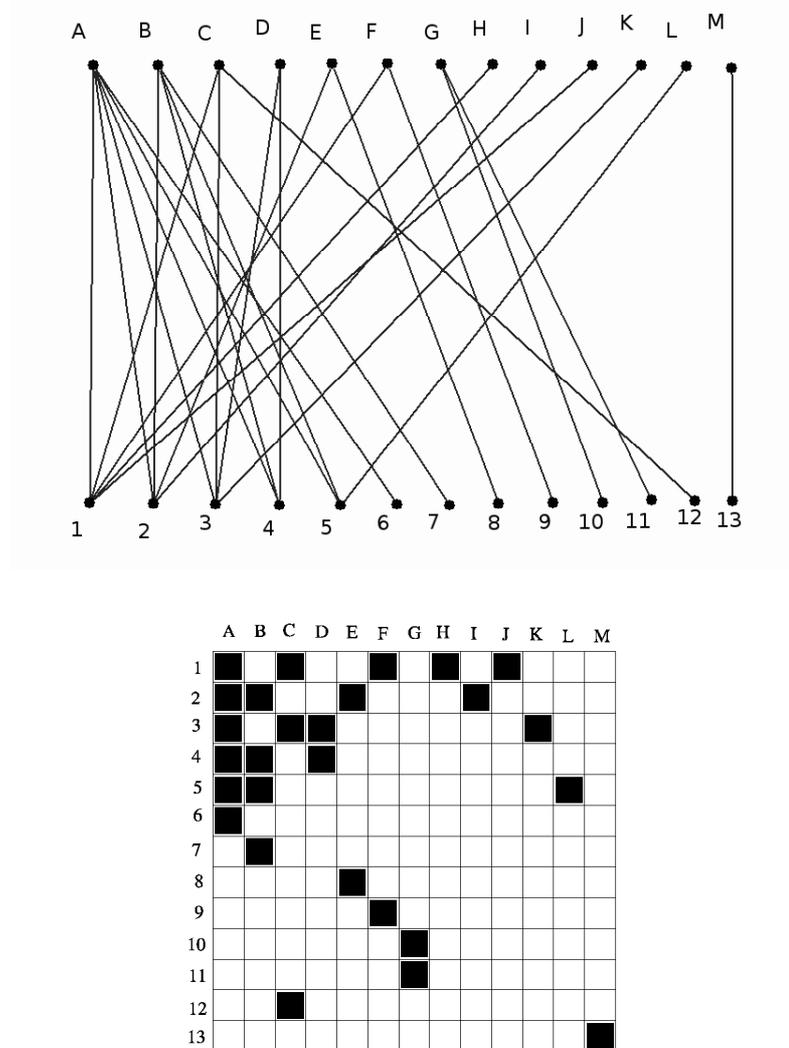


Figura 1.1: Grafo bipartido representando uma rede de herbivoria com 13 espécies de plantas (A a M) e 13 espécies de insetos endófagos (1 a 13). Cada vez que um inseto emerge de uma flor uma conexão entre ambos é estabelecida e representada por uma linha. Em outras palavras, este inseto particular utiliza esta planta particular como local de alimentação. As informações são sintetizadas pela da matriz de adjacência. Dados coletados em uma região montanhosa do Sudeste do Brasil [1].

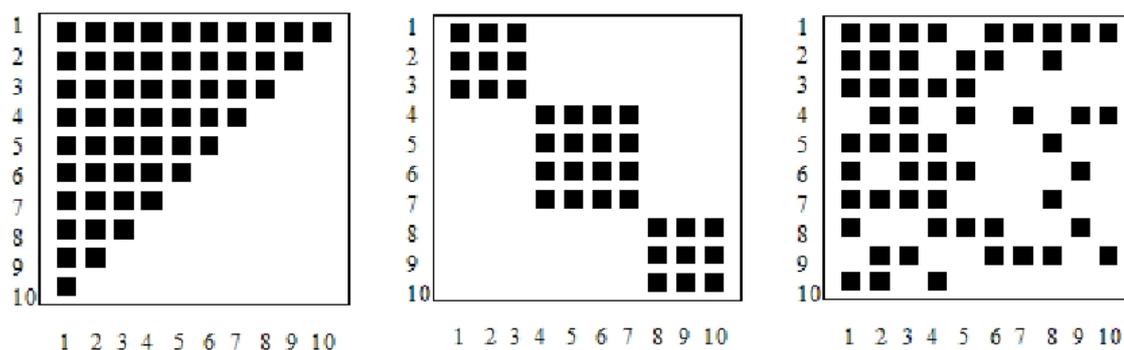


Figura 1.2: Estruturas mais comuns em matrizes. Matriz aninhada à esquerda. Note que os insetos mais generalistas (primeiras colunas) interagem com as plantas mais generalistas (primeiras linhas). Matriz compartimentada no centro. Espécies dentro de um compartimento interagem mais fortemente entre si que com de outros compartimentos. Matriz aleatória à direita. Não há uma estrutura bem definida e as espécies interagem ao acaso.

interage com uma distribuição aleatória de espécies de animais (ou plantas) presentes na comunidade (figura 1.2 à direita) [27].

Os últimos anos mostraram um número crescente de estudos sobre redes, não somente biológicas. Estas redes podem ser formadas por exemplo por páginas da internet e seus links, citações de artigos científicos, relações sociais entre indivíduos em uma empresa, atores de cinema, sistema nervoso, etc. Essas redes são consideradas complexas e são conhecidas como mundos pequenos devido ao pequeno número de ligações entre quaisquer dois nós. Estas teias apresentam a seguinte peculiaridade: cada novo nó (ou espécie) irá se ligar preferencialmente a nós preexistentes com uma probabilidade proporcional ao número de interações desta (quanto mais conectado o nó, maior a chance de um novo nó se ligar a ele), descrevendo uma função do tipo lei de potência ou "power law" e apresentando uma estrutura independente de escala [33, 34]. Teias com distribuições descritas pela função lei de potência são muito sensíveis à perda de nós muito conectados, mais raros, enquanto são bastante robustas à perda aleatória de nós, já que a maior parte destes é pouco conectada [33, 35, 28, 34].

Teias descritas por uma função *power law* têm uma grande chance de apresentar estrutura aninhada, como veremos na próxima seção. A adequação de teias tróficas à função *power law* tem sido testada recentemente e é controversa. Enquanto alguns estudos descreveram teias tróficas como se adequando perfeitamente a esta função [36, 27], ao menos um estudo não encontrou tal relação para um conjunto de várias teias [28].

## 1.2 Teias Bipartidas de Interações Mutualísticas

Bascompte e colaboradores [27] empreenderam uma análise comparativa de 52 teias bipartidas de interações plantas-mutualistas (polinizadores ou dispersores e suas hospedeiras), e mostraram que a grande maioria das teias apresentou um padrão aninhado. Um outro estudo feito com apenas duas teias bipartidas de polinizador-hospedeira encontrou uma estrutura compartimentada [31]. Entretanto, a demonstração de compartimentação é complicada, e o método usado [37] não pode ser considerado um teste estatístico conclusivo [31]. Em um outro estudo com 53 teias planta-polinizador e planta-dispersor, Jordano e colaboradores [38] analisaram a distribuição cumulativa do número de interações por espécie e detectaram um ajuste de praticamente todas as teias ao modelo *power law*, ajuste esse truncado para um grande número de interações. Este ajuste a uma estrutura sem escala, truncado para muitas interações, mostra uma estrutura que parece ser característica para estas teias de mutualistas: um grande número de espécies especialistas coexistem com um número relativamente pequeno de espécies supergeneralistas. A adequação das teias bipartidas de mutualistas tanto ao modelo *power law* quanto ao modelo aninhado mostra como estas duas estruturas estão próximas.

## 1.3 Teias Bipartidas de Interações Herbívoro-Hospedeira

Espécies de herbívoros estão distribuídas em um gradiente que vai desde espécies que se alimentam de uma única espécie de planta hospedeira até espécies que se alimentam de uma ampla gama de hospedeiras de várias famílias. Elas são divididas normalmente em monófagas, oligófagas e polífagas. Espécies monófagas se alimentam de uma única espécie de hospedeira (ou de espécies de um mesmo gênero), enquanto espécies oligófagas se alimentam de vários gêneros dentro de uma mesma família e as polífagas se alimentam de várias famílias diferentes de hospedeira. Das espécies de insetos conhecidas, aproximadamente 75% são monófagas ou oligófagas; e dentre as maiores ordens (Hemiptera, Diptera, Himenoptera e Lepidoptera) mais de 80% das espécies são monófagas ou oligófagas [39].

Herbívoros podem ser divididos em três classes de acordo com o tipo de interação com a hospedeira: parasitas, pastadores e predadores [40]. Os parasitas são espécies que completam um estágio de desenvolvimento dentro de um único indivíduo hospedeiro, causando diminuição da aptidão ou *fitness* deste. Diferente de organismos de vida livre, parasitas não têm a opção de deixar seu hospedeiro inicial e escolher entre outros hospedeiros. Ele deve estar preso de forma segura ao hospedeiro, e enfrentar defesas químicas e físicas de sua hospedeira, além de lidar com potenciais predadores e competidores, assim

como parasitóides, sem sair da hospedeira. Nesta classe estão incluídas espécies endófagas (minadores, galhadores, brocadores e endófagos de capítulos florais), assim como vários grupos de insetos fitófagos como coleópteros e lepidópteros. Pastadores são espécies que mudam de indivíduo hospedeiro durante um estágio de desenvolvimento sem necessariamente matá-lo, enquanto predadores rapidamente matam suas presas. Exemplos de pastadores são gafanhotos, enquanto exemplos de predadores são bruquídeos predadores de sementes. Destas três classes, é esperado que os parasitas sejam os mais especialistas, pois estão em contato mais íntimo com a hospedeira [41, 40]. O fato de as outras duas classes necessitarem de mais de um indivíduo durante seu desenvolvimento favorece a seleção de vários hospedeiros, ou generalização.

Restrições fisiológicas, comportamentais ou mesmo morfológicas ao uso de hospedeiras podem fazer com que espécies de animais mais aparentados usem hospedeiras aparentadas, dividindo a comunidade em compartimentos. A comunidade em questão estará compartimentada se existir convergência no uso de plantas hospedeiras, em circunstâncias em que todos os animais em um compartimento respondem às mesmas características das plantas [42]. A busca de compartimentos em teias tróficas é freqüente, e está relacionada à busca de estabilidade em sistemas tróficos [43]. Ainda não há, entretanto um teste estatístico conclusivo para a existência de compartimentos [37, 44, 31]. Estudos publicados variam entre uma simples análise visual da teia, sem detectar compartimentação [21, 45]; um misto de análise visual e modelos nulos, onde se observou compartimentação [44]; e um método desenvolvido por Raffaelli e Hall [37], usado recentemente por Dicks e colaboradores [31], observando compartimentação. Dois estudos recentes ainda não publicados com endófagos de capítulos de Asteraceae testaram a presença de compartimentos através da técnica de MRPP (Multi-Response Permutation Procedure) [46]: *Tomoplagia* (Tephritidae: Diptera) endófagos de *Vernonieae* (Asteraceae) na Serra do Espinhaço apresentaram uma matriz compartimentada [47], enquanto nenhum padrão especial foi observado para Tephritidae endófagos de *Eupatorieae* (Asteraceae) na Serra da Mantiqueira [48].

## 1.4 Redes de Mutualismo e Herbivoria

Apesar de à primeira vista mutualismo parecer ser distinto de herbivoria, as forças atuando sobre ambas as interações são semelhantes: mutualismos também envolvem o uso e manipulação da morfologia, fisiologia e comportamento da outra espécie de forma a aumentar sua aptidão. Entretanto, padrões observados em teias de mutualistas não necessariamente serão observados em teias bipartidas de herbívoros. Espécies com mutualismos

do tipo frugivoria (dispersão de sementes) e polinização são caracterizadas por uma espécie agindo como hospedeira e a outra agindo como visitadora. Extrema especialização é rara neste tipo de interação por duas razões. A primeira é que os recursos explorados por mutualistas visitantes são quase sempre efêmeros e não permitem a especialização a uma única espécie hospedeira. A segunda razão é que mutualismos se diversificam nas comunidades ao longo do tempo evolutivo à medida que novas espécies são adicionadas à comunidade, não permitindo desta forma a especialização a uma única hospedeira. A flexibilidade de dieta encontrada em quase todas as espécies de vertebrados frugívoros e nectarívoros sugere que a seleção para especialização ocorre menos nestes grupos de vida livre que em predadores e pastadores [40]. Alguns autores [31] defendem que seriam esperados compartimentos em teias de polinizadores, causados pelas síndromes de polinização encontradas em plantas zoófilas. Plantas polinizadas por abelhas possuem flores com muito néctar e corolas longas, enquanto plantas polinizadas por besouros ou dípteros possuem flores pequenas, abertas e com muito pólen, o que poderia levar à subdivisão da comunidade em compartimentos. Muitas interações não observadas em teias tróficas são ditas interações proibidas [38]. Muitas espécies não se encontram associadas em uma comunidade devido a incompatibilidades fenológicas, por exemplo. Mas as interações proibidas não irão ocorrer somente em teias de mutualistas. Herbívoros estão restritos a plantas das quais são capazes de se alimentar, ultrapassando assim barreiras físicas, como ceras e pelos, e químicas, como diferentes compostos secundários. Em teias de mutualistas (polinizadores e dispersores) a maioria das espécies é de vertebrados e lepidópteros, enquanto nas teias de herbívoros a maioria das espécies é de insetos. Animais envolvidos nas teias de mutualistas são em sua grande maioria bem estudados e conhecidos, podendo ser identificados no campo. Uma parte significativa das espécies de insetos herbívoros, no entanto, pode apresentar problemas em sua identificação. A maioria dos insetos herbívoros presentes em teias publicadas estará incluída no grupo de parasitas, indicando uma associação mais íntima com suas hospedeiras que os dispersores e polinizadores. Pastadores são em sua maioria oligófagos e é esperado, ao menos para este grupo uma conformação de teias na forma de compartimentos.

## 1.5 Teias Bipartidas em Biogeografia e Ecologia de Conservação

Biogeografia é o estudo da distribuição da biodiversidade no espaço e no tempo. Seu objetivo de estudo é a localização e a abundância dos organismos, o que é feito levando em conta fatores históricos como especialização, extinção, deslocamentos continentais,

glaciação, variação no nível do mar ou desvios de rios, em conjunto com o isolamento de regiões e suprimentos de energia disponíveis [49].

O conceito de redes bipartidas é muito utilizado em uma sub área da biogeografia chamada biogeografia de ilhas, que tem como objetivo estudar a riqueza de espécies em comunidades naturais. Além de ilhas propriamente ditas, podem ser consideradas como tal montanhas cercadas por desertos, lagos cercados por terra firme ou fragmentos de florestas em meio a paisagens alteradas por fatores antrópicos, por exemplo. Atualmente o termo é usado para referir-se a qualquer ecossistema envolto por ecossistemas diferentes. Esta teoria propõe que o número de espécies encontrado em uma ilha não perturbada é determinada por imigração, emigração e extinção [50]. Tanto a imigração como a emigração são afetadas pela distância de uma ilha a uma fonte de colonizadores, de modo que ilhas mais isoladas tendem a apresentar um menor grau desses dois fenômenos que as menos isoladas. Normalmente a referida fonte de colonização é o continente, mas pode também ser outra ilha.

A taxa de extinção é afetada pelo tamanho da ilha. Ilhas maiores contém habitats mais variados, e por isso reduzem a probabilidade de extinção devida a eventos esporádicos, uma vez que apresentam mais recursos alternativos. O sucesso da imigração em habitats heterogêneos alcançará portanto um maior número de espécies em relação a habitats menores e mais homogêneos. O equilíbrio dinâmico entre extinção e emigração por um lado e imigração por outro deve assegurar a manutenção do nível de riqueza de espécies.

O aninhamento nestes casos significa que espécies presentes em fragmentos ou ilhas menores são subconjuntos das espécies encontradas em fragmentos sucessivamente maiores, o que significa uma vulnerabilidade diferencial à extinção que pode ser previsível, tendo uma relação funcional com a dimensão do fragmento [66, 52]. Em um padrão perfeitamente aninhado, o fragmento maior terá sempre mais espécies que qualquer combinação possível de fragmentos menores cujas áreas somadas sejam a mesma que a do fragmento maior [51]. Neste contexto, a análise do aninhamento pode adicionar um novo elemento ao debate SLOSS, pois sistemas não aninhados indicam que mesmo pequenos fragmentos são importantes. SLOSS é um acrograma para *Single Large or Several Small*, e refere-se a um debate ocorrido na ecologia durante os anos 70 e 80, que discutia a melhor forma de conservação da biodiversidade em habitats fragmentados: uma única grande reserva ou muitas reservas menores. Conhecer o grau de aninhamento de um sistema biogeográfico pode ajudar a entendê-lo e descrevê-lo, embora não possa ser usado como único fator a ser considerado. Segundo Whittaker e Fernández-Palacios [52], "um índice de aninhamento pode realizar uma descrição composicional e talvez contribua para

a identificação de espécies inclinadas à extinção, mas à sua análise não deve ser dada prioridade em se tratando de planos de conservação”.

## 2 *Conceitos Matemáticos*

### 2.1 Grafos

A Teoria de Grafos é amplamente utilizada em modelos matemáticos e constitui o modo mais direto de descrição de situações onde o conceito de redes é empregado. Neste capítulo abordamos sucintamente as definições básicas desta teoria, sua descrição matricial e também alguns conceitos métricos utilizados nesta descrição. A abordagem feita aqui baseia-se em sua maior parte no livro de Gary Chartrand [53]. Um grafo  $G$  é um conjunto não vazio  $V$  onde define-se uma relação  $R$  simétrica e irreflexiva. Sendo  $R$  simétrica, para cada par ordenado  $(u, v) \in R$ , o par  $(v, u)$  também pertence a  $R$ . Definimos  $E$  como o conjunto de pares simétricos em  $R$ . Por exemplo, um grafo pode ser definido pelo conjunto

$$V = \{v_1, v_2, v_3, v_4\}$$

Juntamente com a relação

$$R = \{(v_1, v_2), (v_1, v_3), (v_2, v_1), (v_2, v_3), (v_3, v_1), (v_3, v_2), (v_3, v_4), (v_4, v_3)\}.$$

Nesse caso,

$$E = \{\{(v_1, v_2), (v_2, v_1)\}, \{(v_1, v_3), (v_3, v_1)\}, \{(v_2, v_3), (v_3, v_2)\}, \{(v_3, v_4), (v_4, v_3)\}\}.$$

No grafo  $G$ , cada elemento de  $V$  é denominado vértice e  $V$  é o conjunto vértice. O número de vértices de  $G$  é chamado ordem de  $G$ . Cada elemento de  $E$  (isto é, cada conjunto consistindo em dois pares ordenados simétricos em  $R$ ) é chamado de aresta e  $E$  é chamado de conjunto aresta de  $G$ . Por isso,  $|V| =$  ordem de  $G$  e  $|E| =$  tamanho de  $G$ .

Se  $G$  é um grafo definido em termos de um conjunto vértice  $V$  e uma relação  $R$  em  $V$ , então  $(u, v) \in R$  implica em  $(v, u)$  pertence a  $R$ . Por isso,  $(u, v), (v, u)$  é uma aresta de  $G$ . É conveniente designar tal aresta por  $uv$  (ou, de forma equivalente,  $vu$ ). O conjunto aresta  $E$  determina completamente a relação  $R$ . De fato, costuma-se descrever um grafo em termos de seu conjunto vértice e de seu conjunto aresta. O grafo  $G$  ilustrado acima pode então ser definido pelos conjuntos  $V = \{v_1, v_2, v_3, v_4\}$  e  $E = \{v_1v_2, v_1v_3, v_2v_3, v_3v_4\}$ .

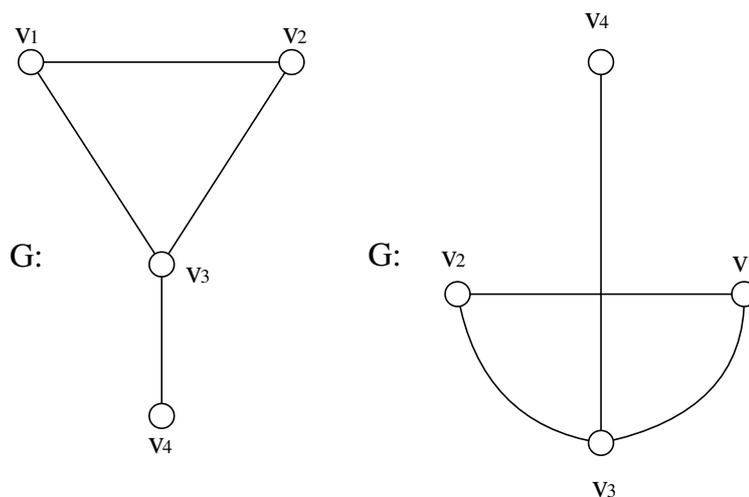


Figura 2.1: Diagramas representando grafos

Portanto, a ordem de  $G$  é quatro, assim como seu tamanho. O conjunto vértice e o conjunto aresta de  $G$  são expressos por  $V(G)$  e  $E(G)$ , respectivamente

Uma vez que o subconjunto vazio de  $V \times V$  é uma relação irreflexiva e simétrica em  $V$ , isso resulta que o conjunto aresta do grafo pode ser vazio, ou seja, o grafo pode não ter arestas, mas por definição tem que possuir vértices.

É frequentemente útil representarmos grafos por meio de diagramas, onde indicamos os vértices por pontos ou pequenos círculos e representamos as arestas por segmentos de linha ou curvas unindo os pontos. Os segmentos de linha ou curvas são desenhados de modo que passem apenas pelos dois pontos que unem. Diagramas representando o grafo  $G$  descrito acima são mostrados na figura 2.1. O primeiro usa apenas segmentos retos, enquanto o segundo emprega linhas curvas. Apesar de os diagramas parecerem diferentes eles contém exatamente os mesmos vértices e as mesmas arestas, e assim descrevem o mesmo grafo. Note que no segundo diagrama os segmentos de linha que representam as arestas  $v_1v_2$  e  $v_3v_4$  se interceptam. Isso pode acontecer (de fato, pode ser inevitável), mas esse ponto de intersecção não é um vértice. Como mencionado anteriormente temos quatro vértices para este exemplo.

A partir de um diagrama como o mostrado na figura 2.1 podemos descrever completamente o grafo e é conveniente referir-se ao diagrama do grafo  $G$  como sendo o próprio grafo  $G$ . Seguem agora algumas definições elementares inspiradas nos aspectos geométricos dos grafos.

Se  $e = uv \in E(G)$  (ou seja,  $uv$  é uma aresta do grafo  $G$ ) dizemos que  $e$  une-se aos vértices  $u$  e  $v$ . Os vértices  $u$  e  $v$  são adjacentes no grafo  $G$  se  $uv \in E(G)$ . Dizemos que  $u$

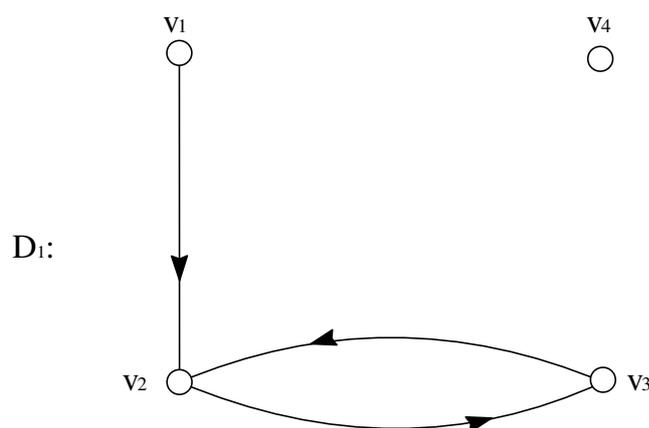


Figura 2.2: Grafo orientado ou dígrafo

e  $v$  são adjacentes um ao outro. Se  $uv \notin E(G)$  então  $u$  e  $v$  serão vértices não adjacentes. Se  $e = uv \in E(G)$  então  $u$  e  $v$  são ditos incidentes em  $e$ . Se  $uv$  e  $uw$  são arestas distintas do grafo  $G$  ( $v \neq w$ ), então  $uv$  e  $uw$  são arestas adjacentes. Por isso  $v_1$  e  $v_3$  são adjacentes e  $v_1$  e  $v_4$  não são adjacentes no grafo  $G$  da figura 2.1. O vértice  $v_3$  é incidente à aresta  $v_2v_3$ , mas  $v_4$  não é incidente em  $v_2v_3$ . As arestas  $v_1v_3$  e  $v_3v_4$  são adjacentes, mas  $v_1v_2$  e  $v_3v_4$  não são adjacentes.

### 2.1.1 Grafos orientados como modelos matemáticos

Um grafo orientado  $D$  (também chamado dígrafo) é formado por um conjunto finito não vazio  $V$  onde atua uma relação irreflexiva  $R$ , como nos grafos. Cada par ordenado em  $R$  é dita uma aresta orientada ou um arco. Vamos denotar a relação por  $E$  em vez de  $R$  ao tratar de dígrafos.

Como a relação que define o dígrafo  $D$  não precisa ser simétrica, se  $(u, v)$  é um arco de  $D$  não implica que  $(v, u)$  seja um arco do dígrafo. Isso é indicado em um diagrama de  $D$  por meio de um segmento de linha reta ou uma curva entre  $u$  e  $v$  acrescida de uma seta que as "direciona". Considerando que  $(u, v)$  e  $(v, u)$  podem ser arcos de  $D$ , podemos desenhar duas curvas entre  $u$  e  $v$  colocando uma ponta de seta em cada curva em direções opostas.

Se definimos  $V_1 = \{v_1, v_2, v_3, v_4\}$  e  $E_1 = \{(v_1, v_2), (v_2, v_3), (v_3, v_2)\}$ , temos o dígrafo  $D_1$  da figura 2.2.

A relação definindo um dígrafo  $D$  pode ser simétrica. Dígrafos simétricos são grafos e a única diferença entre eles é sua representação pictórica, como mostra a figura 2.3.

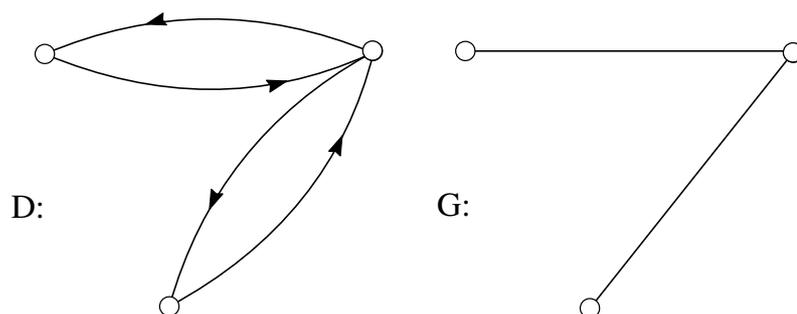


Figura 2.3: Dígrafos simétricos

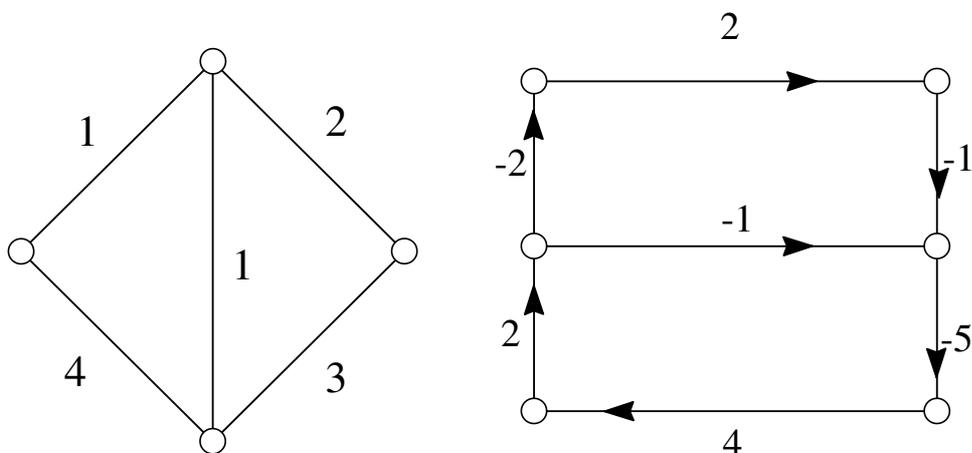


Figura 2.4: Grafos representando redes

### 2.1.2 Redes como modelos matemáticos

Há ocasiões nas quais grafos ou dígrafos não são totalmente apropriados para sua descrição, embora pareçam estar envolvidos. Consideraremos aqui alguns casos deste tipo.

Definimos rede como um grafo ou um dígrafo juntamente com uma função que mapeia o conjunto aresta dentro do conjunto dos números reais. Podemos ter redes não direcionadas como na figura 2.4 à esquerda ou redes que resultam de dígrafos como na figura 2.4 à direita.

Um grafo assinalado  $S$  é uma rede não direcionada cujos valores funcionais são  $+1$  ou  $-1$ . Referimo-nos a cada aresta do grafo assinalado como sendo positiva ou negativa. Por exemplo, se

$$V = \{v_1, v_2, v_3, v_4\},$$

$$E = \{(v_1, v_2), (v_2, v_3), (v_3, v_2)\}$$

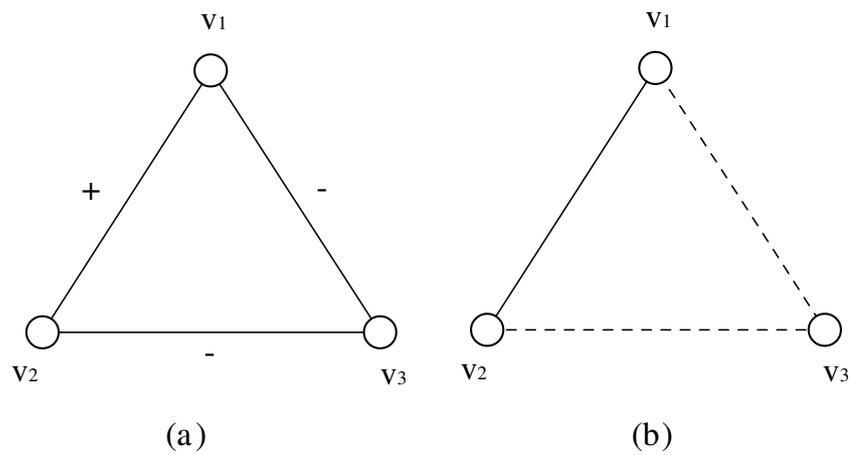


Figura 2.5: Grafos assinalados

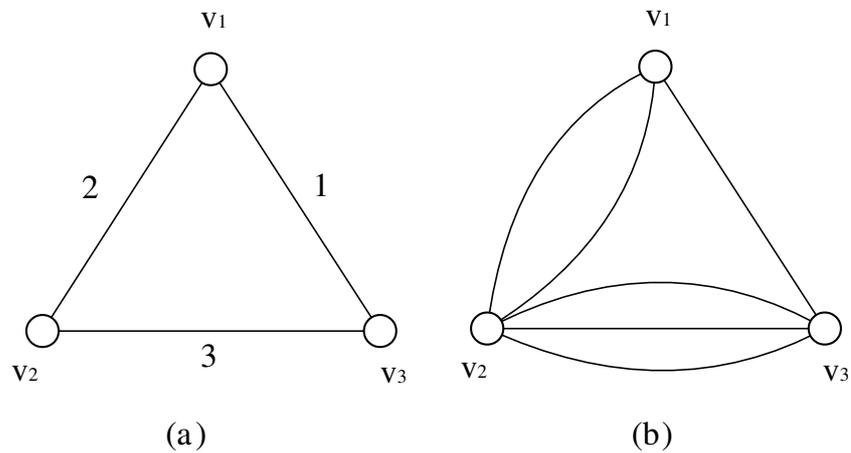


Figura 2.6: Grafos com valores funcionais inteiros

e

$$f = (v_1v_2, +1), (v_1v_3, -1), (v_2v_3, -1),$$

o grafo resultante pode ser representado por uma das duas formas da figura 2.5.

Podemos ter ainda redes cujos valores funcionais são inteiros positivos. Por exemplo, se

$$V = \{v_1, v_2, v_3, v_4\},$$

$$E = \{(v_1, v_2), (v_2, v_3), (v_3, v_2)\}$$

e

$$f = (v_1v_2, 2), (v_1v_3, 1), (v_2v_3, 3),$$

a rede resultante pode ser representada por uma das duas formas da figura 2.6.

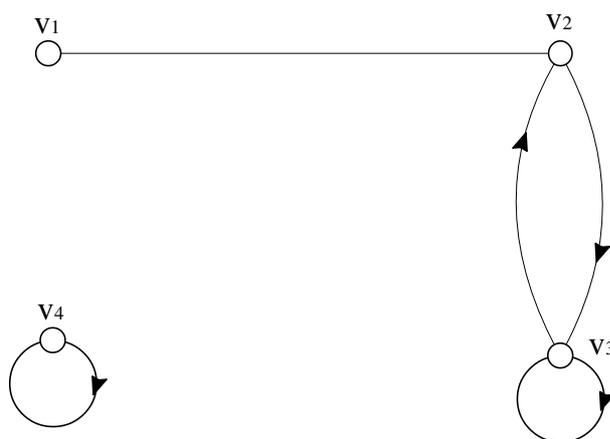


Figura 2.7: Loop-dígrafo

Se uma rede não direcionada é representada por um conjunto de pontos e estes são unidos por um número inteiro de curvas ou segmentos de linha, a rede será chamada de multigrafo. Seja  $M$  um multigrafo com conjunto aresta  $E$  e função associada  $f$ . Se  $uv \in E$  e  $f(uv) = n$ , onde  $n$  é um inteiro positivo, então dizemos que  $u$  e  $v$  são unidos por  $n$  arestas, e referimo-nos a estas arestas como arestas múltiplas.

Em todas as relações que consideramos nesta seção assumimos a irreflexividade. É bem possível que na situação em discussão a relação não seja irreflexiva. Nesse caso, referimo-nos ao par ordenado  $(u, u)$  como um loop. Se removemos a restrição "irreflexiva" da definição de grafo, chamamos o objeto descrito de loop-grafo. Loop-dígrafo, loop-rede e loop-multigrafo são definidos analogamente. Um loop-multigrafo é também denominado pseudografo.

Seja

$$V = \{v_1, v_2, v_3, v_4\}$$

e

$$E = \{(v_1, v_2), (v_2, v_3), (v_3, v_2), (v_3, v_3), (v_4, v_4)\}$$

Temos então um loop-dígrafo que pode ser representado como na figura 2.7.

### 2.1.3 O grau de um vértice

Seja  $v$  um vértice de  $G$ . O número de arestas de  $G$  incidentes em  $v$  é chamado grau de  $v$  em  $G$ . O grau de  $v$  é definido por  $\deg_G v$ , ou simplesmente  $\deg v$ . Para o grafo  $G$  na figura 2.8,  $\deg v_1 = 1$ ,  $\deg v_2 = 2$ ,  $\deg v_3 = 3$ ,  $\deg v_4 = 2$ ,  $\deg v_5 = 0$ .

Um grafo  $(p, q)$  tem ordem  $p$  e tamanho  $q$ . O grafo  $G$  na figura 2.8 é um grafo  $(5, 4)$ . Observamos que a soma dos graus de seus vértices é 8, o que equivale a  $2q$ , como vemos

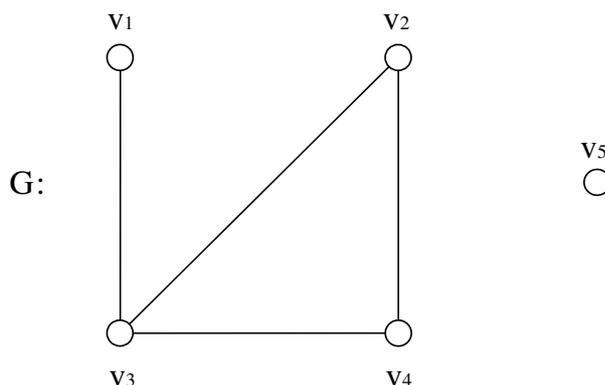


Figura 2.8: Vértices com diversos graus

no teorema 2.1.

### **Teorema 2.1**

Para qualquer grafo  $G$ , a soma de graus de seus vértices equivale a duas vezes o número de arestas em  $G$ . Simbolicamente, se  $G$  é um grafo  $(p, q)$  com vértices  $v_1, v_2, \dots, v_p$ , então

$$\sum_{i=1}^p \deg v_i = 2q.$$

Um vértice é chamado par ou ímpar de acordo com o número do seu grau. O grafo  $G$  na figura 2.8 tem dois vértices ímpares e três vértices pares. O resultado que segue é um corolário do teorema 2.1.

### **Teorema 2.2**

Todo grafo contém um número par de vértices ímpares.

Se todos os vértices do grafo  $G$  têm o mesmo grau  $r$ , dizemos que  $G$  é regular de grau  $r$  ou  $r$ -regular. Um grafo é completo se todos os seus vértices são adjacentes aos demais. Um grafo completo de ordem  $p$  é  $(p-1)$ -regular e é definido por  $K_p$ . Cinco grafos completos são mostrados na figura 2.9.

#### **2.1.4 Grafos isomorfos**

Queremos determinar quais condições dois grafos devem satisfazer para serem iguais. Se  $G_1$  e  $G_2$  são dois grafos equivalentes que servem de modelo para duas situações, então existe alguma similaridade básica entre eles.

Dois grafos  $G_1$  e  $G_2$  são iguais se podemos redesenhar um deles de modo que sua forma fique idêntica a do outro. Os grafos  $G_1$  e  $G_2$  na figura 2.10 têm essa propriedade.

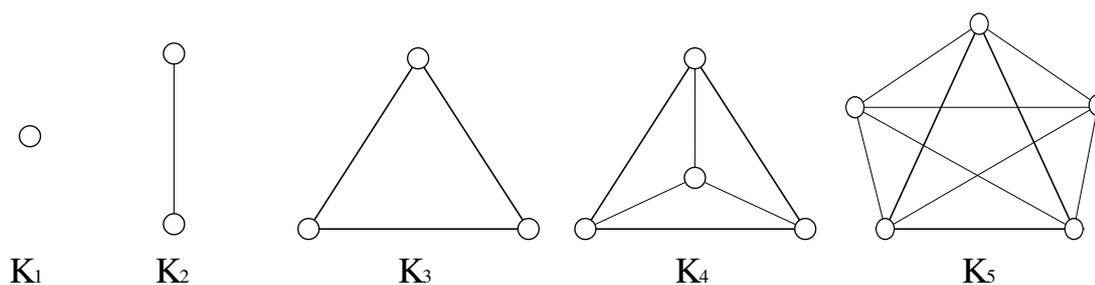


Figura 2.9: Grafos completos

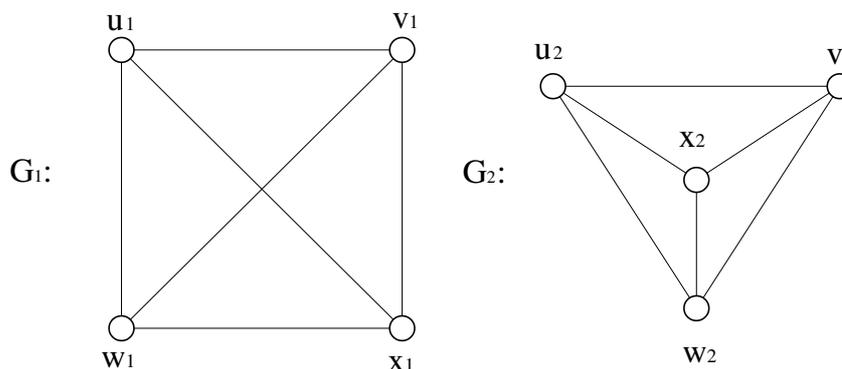


Figura 2.10: Grafos isomorfos

Dois grafos iguais no sentido acima são ditos isomorfos. Um isomorfismo de  $G_1$  em  $G_2$  é um mapeamento unívoco  $\phi : V(G_1) \rightarrow V(G_2)$  tal que dois vértices  $u_1$  e  $v_1$  são adjacentes em  $G_1$  se e somente se os vértices  $\phi(u_1)$  e  $\phi(v_1)$  são adjacentes em  $G_2$ . Dizemos que  $G_1$  e  $G_2$  são isomorfos se existe um isomorfismo de  $G_1$  para  $G_2$ . Se  $\phi$  é um isomorfismo de  $G_1$  em  $G_2$ , o mapeamento inverso  $\phi^{-1}$  de  $G_2$  em  $G_1$  também satisfaz a definição de isomorfismo.

Se  $G_1$  e  $G_2$  são grafos isomorfos existe um mapeamento unívoco  $\phi$  de  $V(G_1)$  para  $V(G_2)$ . Isso implica que  $V(G_1)$  e  $V(G_2)$  têm a mesma ordem. Sejam  $u_1$  e  $v_1$  dois vértices de  $G_1$  e supomos que  $\phi(u_1) = u_2$  e  $\phi(v_1) = v_2$ . Então  $u_1$  e  $v_1$  são adjacentes em  $G_1$  se e somente se  $u_2$  e  $v_2$  forem adjacentes em  $G_2$ . Em outras palavras,  $u_1v_1$  é uma aresta de  $G_1$  se e somente se  $u_2v_2$  for uma aresta de  $G_2$ . Isso implica que  $G_1$  e  $G_2$  têm o mesmo tamanho. Contudo, se dois grafos têm a mesma ordem e o mesmo tamanho, não significa que eles sejam necessariamente isomorfos. Por exemplo, os dois grafos da figura 2.11 têm ordem seis e tamanho nove, mas não são isomórficos.

Pode parecer difícil mostrar que os grafos  $G_1$  e  $G_2$  da figura 2.11 não são isomórficos. Podemos simplificar o problema fazendo observações pertinentes. Considere algum mapeamento unívoco  $\phi$  de  $V(G_1) \rightarrow V(G_2)$ . Os vértices  $v_1, v_2$  e  $v_5$  de  $G_2$  são adjacentes entre

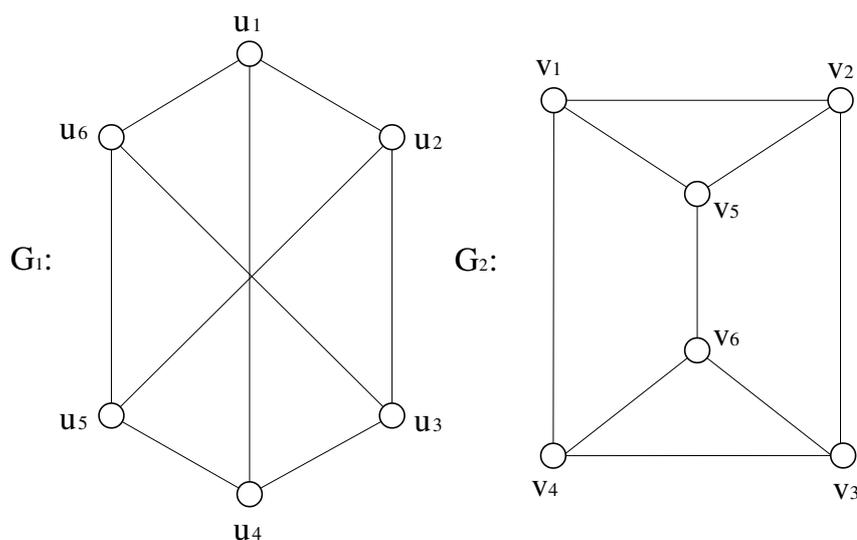


Figura 2.11: Grafos não isomorfos de mesma ordem e tamanho

si, e  $\phi$  deve mapear três vértices de  $G_1$  em  $v_1, v_2$  e  $v_5$ . Se  $\phi$  é um isomorfismo, então dois vértices de  $G_1$  são adjacentes se e somente se os dois vértices imagem da transformação  $\phi$  em  $G_2$  são adjacentes. Isto implica que os três vértices de  $G_1$  cujas imagens são  $v_1, v_2$  e  $v_5$  também precisam ser adjacentes entre si. Entretanto,  $G_1$  não contém três vértices com essa propriedade. Portanto não há isomorfismo de  $V(G_1)$  para  $V(G_2)$  e  $V(G_1)$  não é isomorfo a  $V(G_2)$ .

O teorema a seguir apresenta uma condição necessária mas não suficiente para que dois grafos sejam isomorfos.

### ***Teorema 2.3***

*Se  $G_1$  e  $G_2$  são grafos isomórficos, então os graus dos vértices de  $G_1$  são exatamente os mesmos graus dos vértices de  $G_2$ .*

No entanto, os vértices de dois grafos podem ter exatamente os mesmos graus mas eles podem não ser isomorfos, como  $G_1$  e  $G_2$  da figura 2.11.

Existe somente um grafo de ordem um, tendo necessariamente tamanho zero. Da mesma forma, há apenas um grafo de ordem dois e tamanho zero, e somente um grafo de ordem dois e tamanho um. No entanto, há três grafos de ordem quatro e tamanho três, mostrados na figura 2.12. Podemos afirmar isso de outro modo: entre os grafos de ordem quatro e tamanho três, existem três classes de isomorfismos. Assim, se temos quatro ou mais grafos de ordem quatro e tamanho três, dois ou mais desses grafos devem pertencer à mesma classe.

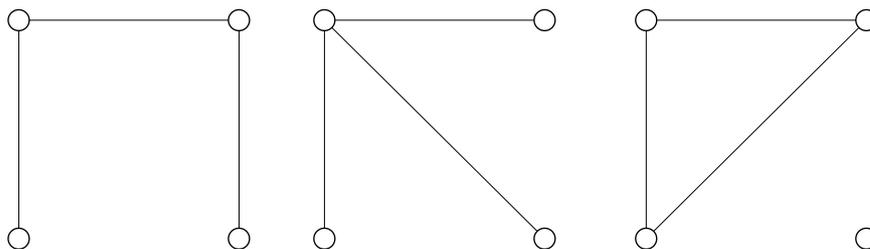
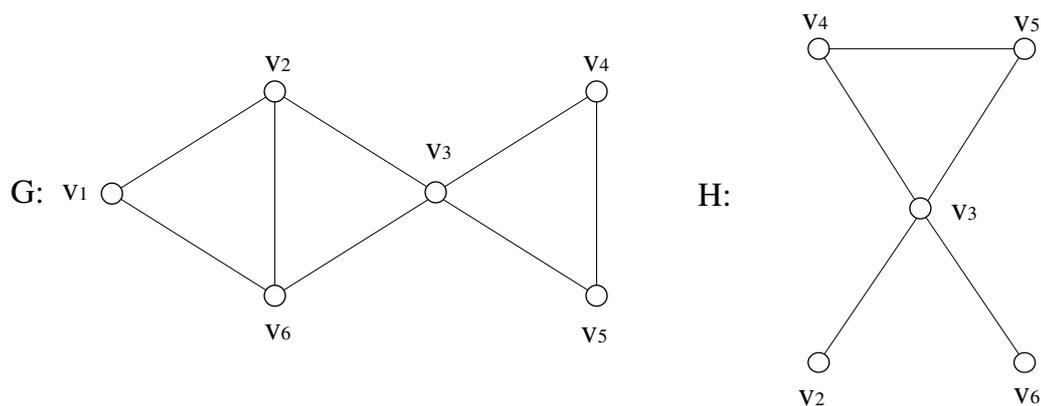


Figura 2.12: Classes de isomorfismo

Figura 2.13: Subgrafo  $H$  conectado ao grafo  $G$ 

### Teorema 2.4

*Princípio da Casa de Pombo (ou Caixas de Dirichlet):* Seja  $S$  um conjunto finito de  $n$  elementos, e seja  $S_1, S_2, \dots, S_k$  uma partição de  $S$  em  $k$  subconjuntos. Então em pelo menos um subconjunto  $S_i$ ,  $1 \leq i \leq k$ , contém pelo menos  $\lceil n/k \rceil$  elementos.

Portanto, se há três classes de equivalência de grafos  $(4,3)$  ( $k = 3$ ) e temos quatro grafos  $(4,3)$  ( $n = 4$ ), deve existir pelo menos  $\lceil 4/3 \rceil = 2$  grafos na mesma classe de equivalência (onde  $\lceil x \rceil$  representa o menor inteiro maior ou igual a  $x$ ).

### 2.1.5 Grafos Conectados

Seja  $G$  um grafo. Um grafo  $H$  é um subgrafo de  $G$  se  $V(H) \subseteq V(G)$  e  $E(H) \subseteq E(G)$ . Se um grafo  $F$  é isomorfo a um subgrafo  $H$  de  $G$  então  $F$  é também chamado subgrafo de  $G$ . A figura 2.13 mostra um grafo  $G$  e um subgrafo  $H$ .

Sejam  $u$  e  $v$  vértices de um grafo  $G$ . Um caminho  $u - v$  em  $G$  é uma sequência alternativa de vértices e arestas de  $G$ , iniciando com  $u$  e terminando com  $v$ , tal que toda aresta une os vértices imediatamente anterior e posterior a ele. Por exemplo,  $v_3, v_3v_2, v_2, v_2v_6, v_6, v_6v_3, v_3, v_3v_4, v_4, v_4v_5, v_5, v_5v_4, v_4$  é um caminho  $v_3v_4$  no grafo  $G$  da

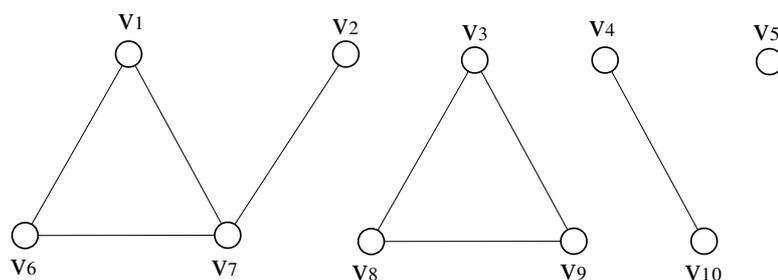


Figura 2.14: Grafo com quatro componentes

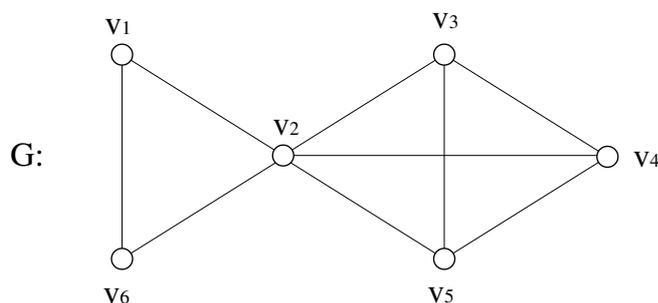


Figura 2.15: Ciclos e circuitos

figura 2.13. Este caminho pode ser expresso mais facilmente como  $v_3, v_2, v_6, v_3, v_4, v_5, v_4$ . Uma trilha  $u - v$  em um grafo é um caminho  $u - v$  que não repete nenhuma aresta. O caminho  $v_3v_4$  descrito acima não é uma trilha  $v_3v_4$ ; no entanto,  $v_3, v_2, v_6, v_3, v_4$  é uma trilha  $v_3v_4$  no grafo  $G$  da figura 2.13.

Uma trajetória  $u - v$  é um caminho  $u - v$  (ou trilha  $u - v$ ) que não repete nenhum vértice. Novamente, no grafo  $G$  da figura 2.13  $v_3, v_5, v_4$  é uma trajetória  $v_3, v_4$ .

Dois vértices  $u$  e  $v$  no grafo  $G$  estão conectados se  $u = v$ , ou se  $u \neq v$  e existe uma trajetória  $u - v$  em  $G$ . Um grafo  $G$  está conectado se toda dupla de vértices de  $G$  está conectada, do contrário  $G$  está desconectado.

Um subgrafo conectado  $H$  do grafo  $G$  é chamado componente de  $G$  se  $H$  não está contido em nenhum subgrafo conectado de  $G$  tendo mais vértices ou arestas que  $H$ . Por exemplo, o grafo da figura 2.14 tem quatro componentes. Se um grafo tem apenas um componente, o grafo está conectado.

Uma trilha  $u - v$  na qual  $u = v$  que contenha pelo menos três arestas é chamada de circuito. Em outras palavras, um circuito deve terminar no mesmo vértice em que começou. Um circuito que não repete nenhum vértice (exceto o primeiro e o último) é chamado de ciclo. Por exemplo, no grafo  $G$  da figura 2.15,  $v_1, v_2, v_3, v_5, v_2, v_6$  é um circuito mas não é um ciclo, enquanto que  $v_2, v_4, v_3, v_5, v_2$  é um ciclo (e também um circuito).

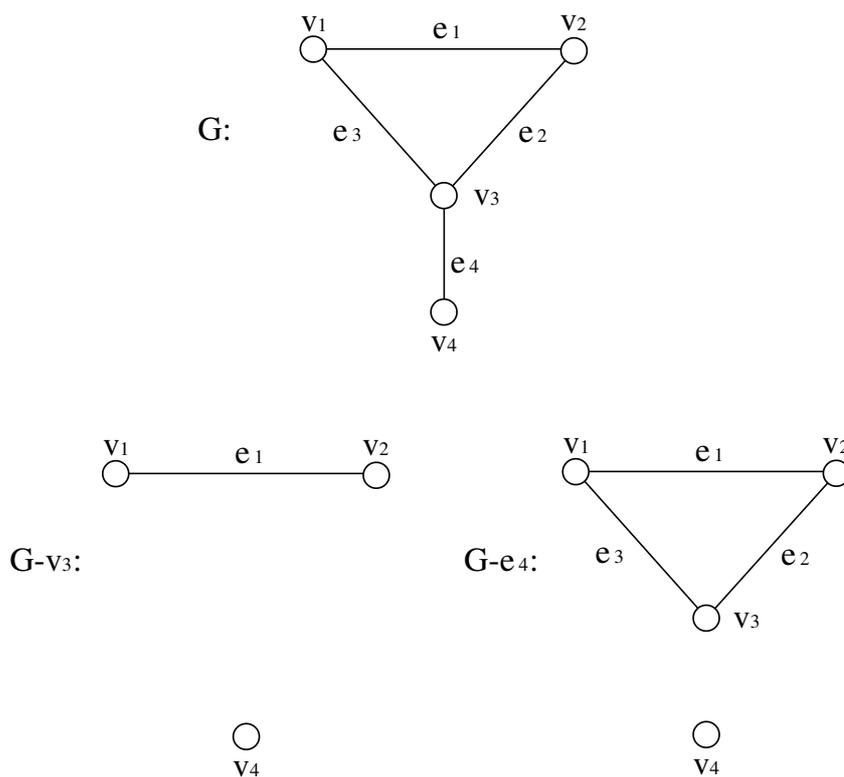


Figura 2.16: Vértices cortados e pontes

Pela definição, uma trilha é uma seqüência alternativa de vértices e arestas, embora concordemos em representar uma trilha por uma seqüência de vértices. Os conjuntos de vértices e de arestas determinados pela trilha produzem um subgrafo. Em geral, consideramos o subgrafo formado por vértices e arestas de um objeto (trilha, trajetória, circuito ou ciclo) como sendo o próprio objeto.

### 2.1.6 Vértices Cortados e Pontes

Se  $e$  é uma aresta do grafo  $G$ , então  $G - e$  é um subgrafo de  $G$  com o mesmo conjunto vértice de  $G$  com todas as arestas de  $G$  exceto  $e$ . Se  $v$  é um vértice de um grafo  $G$  que contém pelo menos dois vértices, então  $G - v$  é um subgrafo de  $G$  cujo conjunto vértice consiste de todos os vértices de  $G$  exceto  $v$  e cujo conjunto aresta consiste em todas as arestas de  $G$  exceto aquelas incidentes em  $v$ . A figura 2.16 ilustra esses conceitos.

Um vértice  $v$  em um grafo conectado  $G$  é chamado de vértice cortado se  $G - v$  está desconectado, como o vértice  $v_3$  no grafo da figura 2.16. Uma aresta  $e$  em um grafo conectado  $G$  é chamada de ponte se  $G - e$  é desconectado, como a aresta  $e_4$  do grafo  $G$  na figura 2.16.

Se  $v$  é um vértice cortado de um grafo  $G$  conectado, então  $G - v$  contém dois ou

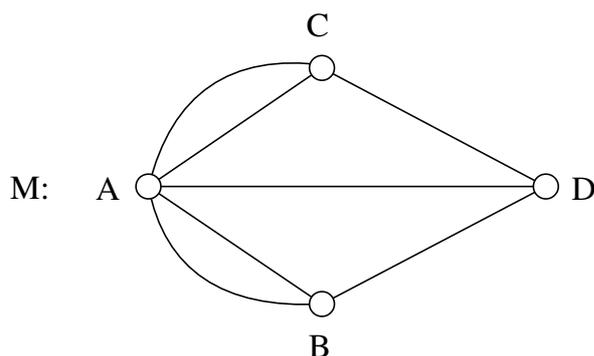


Figura 2.17: Multigrafo representando as pontes de Königsberg

mais componentes. Entretanto, se  $e$  é uma ponte de  $G$  então  $G - e$  tem exatamente dois componentes. O teorema a seguir traz uma condição para que arestas sejam pontes.

### ***Teorema 2.5***

*Seja  $G$  um grafo conectado. Uma aresta  $e$  de  $G$  é uma ponte de  $G$  se e somente se e não aparece em nenhum ciclo de  $G$ .*

## **2.2 O problema da ponte de Königsberg**

O famoso problema da ponte de Königsberg foi resolvido por Euler em 1736 e marca o início da Teoria de Grafos: *"Na cidade de Königsberg havia no século 18 sete pontes sobre o rio Pregel. Elas conectavam duas ilhas que ficavam no rio entre si e com as margens opostas. As pessoas da cidade divertiram-se por muito tempo com o seguinte problema: é possível atravessar as sete pontes em um passeio sem cruzar mais de uma vez alguma delas?"*[54].

A situação em Königsberg pode ser representada convenientemente por um multigrafo, como mostra a figura 2.17. O conjunto vértice representa as regiões de terra firme e cada dupla de vértices é unida por um número de arestas iguais ao número de pontes unindo as correspondentes regiões no solo.

O problema das pontes de Königsberg resume-se essencialmente em determinar se o multigrafo  $M$  da figura 2.17 tem um caminho contendo todas as suas arestas. Por tentativa e erro concluiríamos que não existe este caminho, mas apresentaremos uma prova para isso no teorema abaixo.

**Teorema 3.1**

O multigrafo  $M$  da figura 2.17 não tem uma trilha que contenha todas as arestas de  $M$ .

Prova

Primeiramente, note que este teorema é negativo por natureza, desde que desejamos mostrar que  $M$  não possui esse tipo de caminho. Portanto, é natural tentar uma prova por contradição. Suponha que o multigrafo da figura 2.17 tem um caminho, dito  $P$ , que contém todas as arestas de  $M$ . Então  $P$  começa em um dos quatro vértices  $A$ ,  $B$ ,  $C$  ou  $D$  e termina em um dos vértices  $A$ ,  $B$ ,  $C$  ou  $D$  (o mesmo vértice em que  $P$  começou se  $P$  é um circuito). Temos pelo menos dois vértices entre  $A$ ,  $B$ ,  $C$  e  $D$  tais que  $P$  não começa nem termina neste vértice. Portanto, há pelo menos um vértice entre  $B$ ,  $C$  e  $D$  no qual  $P$  não começa nem termina. Denotamos este vértice por  $v$ .

Note que cada um dos vértices  $B$ ,  $C$  e  $D$  tem grau três. Deste modo, após alguma aresta de  $P$  entrar no vértice  $v$  pela primeira vez e outra aresta de  $P$  deixar o vértice  $v$ , haverá exatamente uma aresta incidente em  $v$  que não pertence ainda a  $P$ . Agora,  $v$  deve ser inserido ao longo do caminho  $P$  mais uma vez por meio da aresta incidente a  $v$  que não foi usada ainda. Entretanto, após chegar pela segunda vez em  $v$ , não encontramos arestas restantes para a saída, de forma que  $P$  encerra-se em  $v$ . Isso é impossível desde que  $P$  não termine em  $v$ . Portanto, não existe o caminho  $P$ , produzindo a contradição desejada.

## 2.3 O problema da dança: Grafos Bipartidos

Suponha que você tenha um grupo de homens e mulheres em uma festa. Em que condições é possível que todas as mulheres dançam ao mesmo tempo?

Exploraremos esse problema configurando um modelo matemático usando grafos. Seja  $G$  um grafo cujo conjunto vértice represente as pessoas que estão na festa, de tal modo que dois vértices são adjacentes se as pessoas a eles correspondentes são parceiros compatíveis na dança. Será possível todas as mulheres dançarem ao mesmo tempo se e somente se  $G$  contiver um subgrafo 1-regular  $F$  tal que o número de arestas em  $F$  é igual ao número de mulheres.

Alguns conceitos podem ser definidos a partir desse problema de fácil discussão. Um grafo  $G$  é chamado bipartido se for possível realizar uma partição do conjunto vértice de  $G$  em dois subconjuntos  $V_1$  e  $V_2$ , de forma que toda aresta de  $G$  pode unir um vértice de  $V_1$  com um vértice de  $V_2$  mas não une vértices do próprio subconjunto. Na figura 2.18 redesenhamos um grafo  $G$  para evidenciar sua propriedade de bipartição. Neste grafo

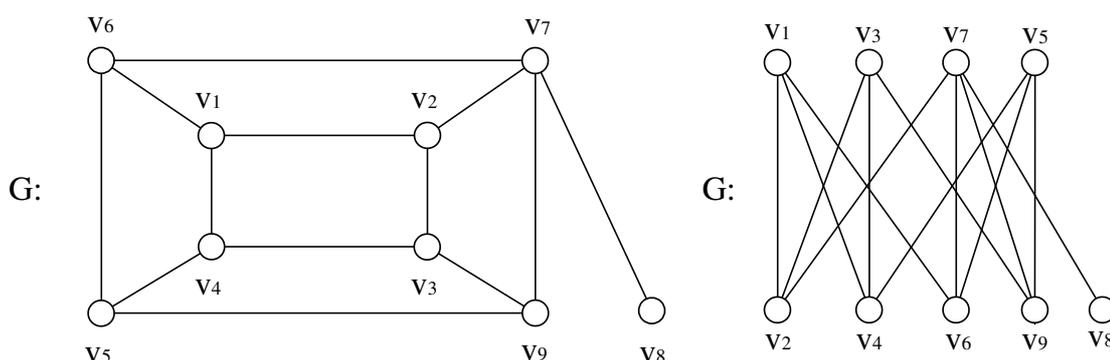


Figura 2.18: Grafos bipartidos

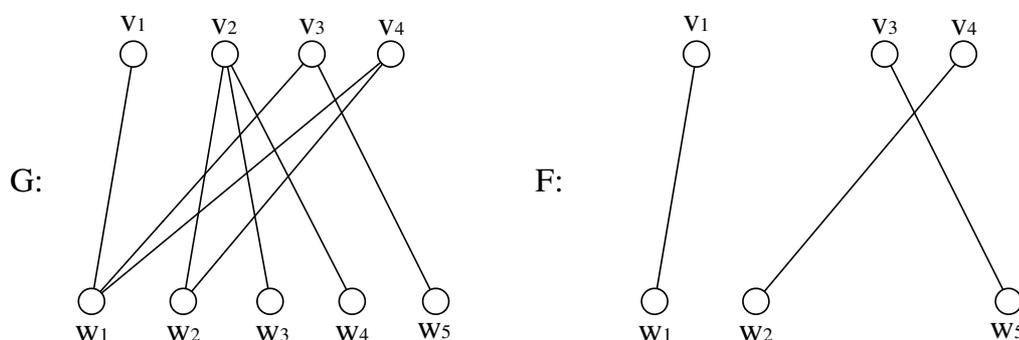


Figura 2.19: Vértices emparelhados

temos  $V_1 = \{v_1, v_3, v_5, v_7\}$  e  $V_2 = \{v_2, v_4, v_6, v_8, v_9\}$ .

Seja  $G$  um grafo bipartido cujo conjunto vértice é dividido nos subconjuntos  $V_1$  e  $V_2$ . Seja  $U_1$  um subconjunto de  $V_1$ . Dizemos que  $U_1$  está emparelhado ao subconjunto  $U_2$  de  $V_2$  se  $G$  contiver um subgrafo  $F$  1-regular cujo conjunto vértice seja  $U_1 \cup U_2$ . Se  $U_1$  está emparelhado a  $U_2$ , devemos ter  $|U_1| = |U_2|$ . O subgrafo  $F$  é denominado uma emparelhamento, pois estabelece um emparelhamento entre os conjuntos de vértices  $U_1$  e  $U_2$ .

Ilustramos esse conceito com o grafo da figura 2.19. Aqui temos  $V_1 = \{v_1, v_2, v_3, v_4\}$  e  $V_2 = \{w_1, w_2, w_3, w_4, w_5\}$ . Se definirmos  $U_1 = \{v_1, v_3, v_4\}$  e  $U_2 = \{w_1, w_2, w_5\}$ , vemos que  $G$  contém o subgrafo  $F$  1-regular com conjunto vértice  $U_1 \cup U_2$ . Portanto,  $U_1$  é emparelhado a  $U_2$ . Note que  $v_2$  pode estar ele próprio emparelhado a um subconjunto de  $V_2$ .

Supondo  $G$  um grafo bipartido com seus conjuntos vértice divididos em  $V_1$  e  $V_2$ . Se  $W_1 \subseteq V_1$ , denotamos por  $W_1^*$  os vértices de  $V_2$  adjacentes a pelo menos um vértice em  $W_1$ . A deficiência  $\text{def}(W_1)$  de  $W_1$  em  $G$  é definida por

$$\text{def}(W_1) = |W_1| - |W_1 *|. \quad (2.1)$$

O conjunto  $U_1 \subseteq V_1$  é dito não deficiente em  $G$  se nenhum subconjunto (não vazio) de  $U_1$  tiver deficiência positiva. Notemos que a deficiência positiva de um subconjunto  $W_1$  de  $U_1$  implica na existência de mais vértices em  $W_1$  do que vértices adjacentes aos elementos em  $W_1$ , o que torna impossível emparelhar  $W_1$  a um subconjunto de  $V_2$ .

Todos os subconjuntos de  $V_1$  no grafo  $G$  da figura 2.19 têm deficiência zero ou negativa, ou seja, nenhum subconjunto de  $V_1$  tem deficiência positiva. Deste modo podemos afirmar que  $V_1$  é não deficiente e que  $V_1$  está emparelhado a um subconjunto de  $V_2$ . Estas duas afirmações não são sinônimas como veremos a seguir.

Voltando ao problema da dança, temos a situação representada por um grafo bipartido  $G$ . Se denotarmos por  $V_1$  os vértices correspondentes às mulheres e por  $V_2$  os que correspondem aos homens, então estamos procurando por condições sob as quais  $V_1$  pode ser emparelhado a um subconjunto de  $V_2$ . O teorema 4.1 fornece tais condições.

### ***Teorema 4.1***

*Seja  $G$  um grafo bipartido cujo conjunto vértice é dividido em conjuntos  $V_1$  e  $V_2$  de modo que toda aresta de  $G$  une um vértice de  $V_1$  com um vértice de  $V_2$ . Então  $V_1$  pode estar emparelhado a um subconjunto de  $V_2$  se e somente se  $V_1$  for não deficiente.*

Aplicando o teorema 4.1 ao problema da dança, podemos determinar se existem parceiros compatíveis para todas as mulheres computando o número total de parceiros compatíveis para cada subconjunto de mulheres. Se o número de parceiros compatíveis para cada subconjunto é pelo menos tão grande quanto o número de mulheres no subconjunto, então os parceiros de dança podem ser encontrados para todas as mulheres em uma dança única.

## **2.4 Grafos e Matrizes**

Um grafo é completamente determinado por seu conjunto vértice e pelo conhecimento de que pares de vértices são adjacentes. Essa mesma informação pode facilmente ser dada por uma matriz. De fato, muito da teoria de grafos pode ser estudado no contexto da teoria de matrizes. Existem vantagens para esta abordagem, visto que matrizes podem servir como entrada de dados para trabalhos computacionais. Por outro lado, existem desvantagens em representar grafos como matrizes, pois perde-se o aspecto visual presente nos grafos, o que pode tornar mais difícil o reconhecimento direto de certas propriedades evidentes numa análise gráfica, como a existência de ciclos, circuitos, trilhas, vértices cortados ou pontes.

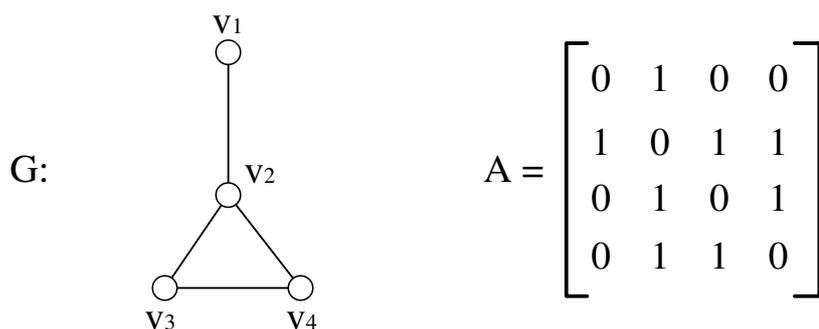
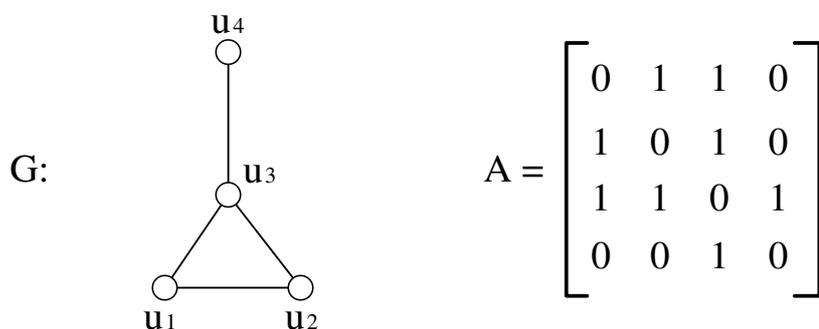
Figura 2.20: Grafo  $G$  e sua matriz de adjacência

Figura 2.21: Grafo da figura anterior com rotulação diferente

Seja  $G$  um grafo de ordem  $p$  com vértices definidos por  $v_1, v_2, \dots, v_p$ . Então a matriz de adjacência  $A = A(G) = [a_{ij}]$  é a matriz  $p \times p$  na qual  $a_{ij} = 1$  se  $v_i$  e  $v_j$  forem adjacentes e  $a_{ij} = 0$  se não o forem. Deste modo a matriz  $A$  é uma matriz binária (ou seja, toda entrada de  $A$  é 0 ou 1). A diagonal principal de  $A$  consiste inteiramente de zeros (ou seja,  $a_{ij} = 0$  para  $i = 1, 2, \dots, p$ ) e  $A$  é simétrica ( $a_{ij} = a_{ji}$  para  $1 \leq i \leq p$  e  $1 \leq j \leq p$ ). Um grafo  $G$  e sua matriz de adjacência são dados na figura 2.20.

Note que a matriz de adjacência de um grafo  $G$  normalmente depende de como rotulamos os vértices. Por exemplo, o grafo  $G$  da figura 2.20 é mostrado novamente na figura 2.21 com uma rotulação diferente, resultando em uma matriz de adjacência diferente. Embora as matrizes das figura 2.20 e 2.21 sejam desiguais, representam grafos isomórficos.

Podemos agora considerar uma segunda matriz que pode ser associada com um grafo. Seja  $G$  um grafo de ordem  $p$  e tamanho  $q$  tal que  $V(G) = \{v_1, v_2, \dots, v_p\}$  e  $E(G) = \{e_1, e_2, \dots, e_q\}$ . A matriz de incidência  $B = B(G) = [b_{ij}]$  é a matriz  $p \times q$  na qual  $[b_{ij}] = 1$  se  $v_i$  é incidente a  $e_j$ , e  $[b_{ij}] = 0$  caso não seja. Da mesma forma que a matriz de adjacência, a matriz de incidência é um matriz  $(0, 1)$ . A matriz de incidência em geral não é simétrica

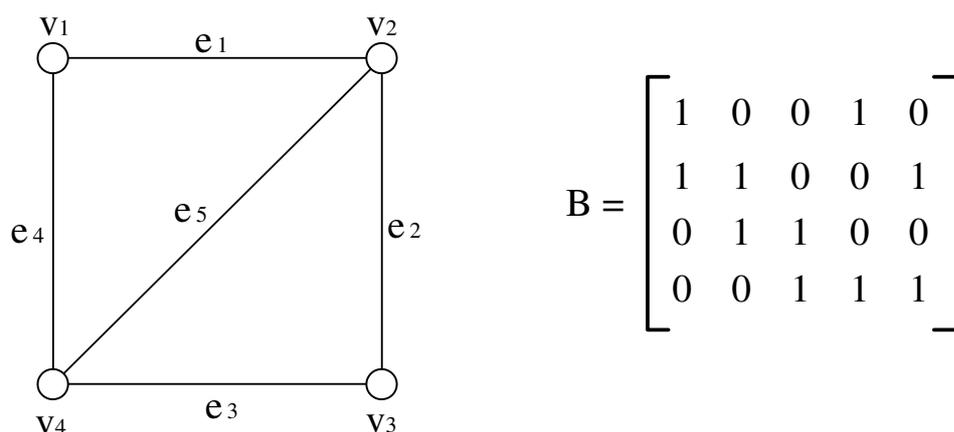


Figura 2.22: Matriz de incidência

e nem quadrada. Uma matriz de incidência particular obtida depende de como rotulamos tanto vértices quanto arestas. A figura 2.22 fornece um exemplo.

Uma matriz de adjacência pode representar também um grafo bipartido. Neste caso os índices  $i$  representam elementos de uma classe e os índices  $j$  representam os elementos de outra classe. Como no caso anterior, os elementos  $a_{ij} = 1$  se  $v_i$  e  $v_j$  forem adjacentes e  $a_{ij} = 0$  se não o forem. A figura 1.1 do capítulo anterior mostra uma matriz deste tipo para interações de herbivoria entre plantas e insetos endófagos.

## 2.5 Conceitos Métricos

Expomos aqui alguns conceitos geométricos que serão utilizados no próximo capítulo em nossas definições de distância entre elementos da matriz de adjacência. Dado um conjunto  $\mathbb{S}$ , uma métrica em  $\mathbb{S}$  é uma função

$$d : \mathbb{S} \times \mathbb{S} \rightarrow \mathbb{R},$$

chamada função distância ou simplesmente distância, que satisfaz as seguintes condições:

1. Não negatividade:

$$d(x, y) \geq 0,$$

2. Identidade dos indiscerníveis:

$$d(x, y) = 0 \iff x = y,$$

3. Desigualdade Triangular:

$$d(x, z) \leq d(x, y) + d(y, z),$$

4. Simetria:

$$d(x, y) = d(y, x),$$

onde os elementos  $x, y$  e  $z$  pertencem ao conjunto  $\mathbb{S}$ .

Exemplos:

No conjunto dos números reais, a métrica usual é dada por:

$$d(x, y) = |x - y|$$

No conjunto  $\mathbb{R}^n$  várias métricas podem ser definidas, por exemplo:

$$d(x, y) = \sqrt[p]{\sum_{i=1}^n |x_i - y_i|^p},$$

$$d(x, y) = \max |x_i - y_i|$$

onde  $p$  é real e positivo.

No conjunto  $\mathbb{C}^0[a, b]$  das funções contínuas no intervalo  $[a, b]$  podemos definir:

$$d(f, g) = \sup_{x \in [a, b]} |f(x) - g(x)|$$

$$d(f, g) = \int_a^b |f(x) - g(x)| dx$$

Em um conjunto  $\mathbb{S}$  qualquer, a métrica discreta é dada por:

$$d(x, y) = \begin{cases} 0, & x = y \\ 1, & x \neq y \end{cases}$$

É possível procedermos a um relaxamento de alguns axiomas métricos, dando margem ao surgimento de várias métricas generalizadas [?]. Como exemplo, se a estrutura que garante a identidade dos indiscerníveis não é satisfeita (os elementos não são distinguíveis) temos uma Pseudométrica. Se a simetria não pode ser obedecida pela função distância temos uma Quasemétrica. Se a distância não apresenta desigualdade triangular estamos falando de uma Semimétrica. No caso de um conjunto não satisfaz solidariamente a identidade dos indiscerníveis, a condição de simetria e a desigualdade triangular temos um caso de Pré-métrica.

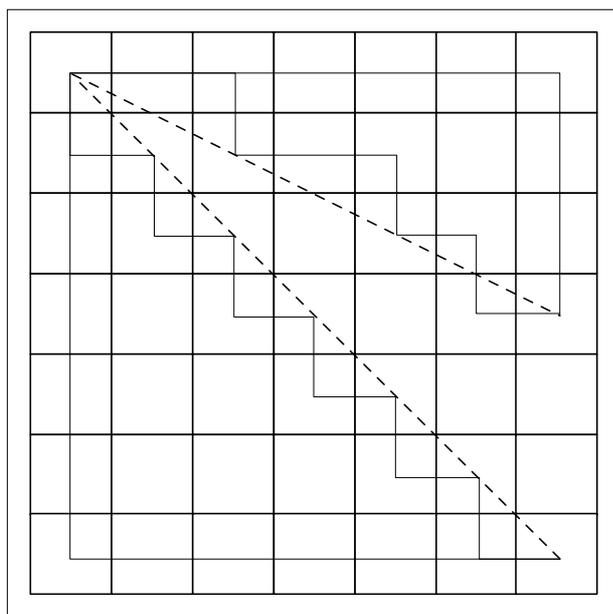


Figura 2.23: Distâncias de Minkowski para  $\chi = 1$ (linhas sólidas) ou  $\chi = 2$ (linhas tracejadas).

## 2.6 Geometria de Manhattan ou *taxicab*

A Geometria de Manhattan ou *taxicab* foi desenvolvida por Hermann Minkowski no século XIX, é uma geometria em que a métrica usual euclidiana é substituída por uma nova métrica em que a distância entre dois pontos é a soma das diferenças absolutas de suas coordenadas. A métrica *taxicab* é também conhecida como distância  $L_1$ , ou distância de Manhattan. O último nome faz alusão ao formato quadriculado da maior parte das ruas na ilha de Manhattan. Tal configuração faz com que a menor distância a ser percorrida por um carro que vai de um ponto a outro na cidade tenha como valor aquele número fornecido pela métrica  $L_1$ . [55] A métrica de Minkowski é tal que a distância de Minkowski de ordem  $\chi$  entre dois pontos  $P = (x_1, x_2, \dots, x_n)$  e  $Q = (y_1, y_2, \dots, y_n) \in \mathbb{R}^n$  é definida por  $(\sum_{i=1}^n |x_i - y_i|^\chi)^{1/\chi}$ . As distâncias de Minkowski são tipicamente usadas com  $\chi = 1$  ou  $\chi = 2$ . No último caso temos a distância Euclidiana e no primeiro temos a distância de Manhattan, como mostra a figura 2.23.

## 3 *Desenvolvimento Analítico do Índice de Aninhamento*

### 3.1 Aninhamento em Matrizes de Adjacência

De um ponto de vista matemático, o objeto central no estudo do aninhamento é uma matriz binária (cujos elementos podem ser zero ou um). Na ecologia de metacomunidades, uma célula ocupada  $a_{i,j}$  significa que a espécie  $j$  está presente no sítio  $i$ , ou que a espécie  $j$  (plantas) interage com a espécie  $i$  (polinizador), por exemplo.

Para visualizar o aninhamento na matriz, é necessário rearranjar as linhas e colunas da mesma. É importante ressaltar que cada vez que uma linha ou coluna é permutada com outra linha ou coluna, as interações entre os elementos dos grupos  $i$  e  $j$  permanecem as mesmas. A matriz binária com que trabalhamos é também chamada matriz de incidência, matriz de presença e ausência, matriz de adjacência na linguagem de teoria de redes ou grade "lattice" no contexto de física da matéria condensada.

Para estabelecer a notação que utilizamos, consideremos uma matriz de tamanho  $L_1$  e  $L_2$ . O número de elementos "1" (célula ocupada) na matriz é  $N$  e deste modo a ocupação, ou fração da matriz ocupada por elementos "1", é  $\rho = \frac{N}{L_1 L_2}$ . Seguindo a argumentação do parágrafo anterior, há  $L_1! \times L_2!$  possibilidades para representar a matriz segundo diferentes permutações de linhas e colunas. Cada uma destas possibilidades nada mais é que a representação da mesma estrutura de rede. O rearranjo de linhas e colunas é uma forma efetiva de visualizar o aninhamento na matriz. Quando rearranjamos os elementos de uma matriz escolhemos uma das  $L_1! \times L_2!$  possibilidades, aquela em que os elementos da matriz estão mais empacotados. Em outras palavras, escolhemos uma representação para a matriz na qual temos um número máximo de elementos não nulos estão o mais perto possível da extremidade da grade onde  $i = 1$  e  $j = 1$ . O processo de empacotamento é um passo usual que precede o cálculo do índice de aninhamento [6, 12].

O problema central que colocamos aqui é como medir o desvio de uma dada matriz de outra que esteja perfeitamente aninhada. Expomos brevemente abaixo três diferentes estimadores presentes na literatura que inspiraram diretamente nossa medida: o estimador

de temperatura de Atmar e Paterson [11], o índice  $N_0$  de Atmar e Paterson [12] e o índice de discrepância de Brualdi and Sanderson [13]. Depois disso introduziremos nosso método que, embora seja baseado nos estimadores prévios, é mais robusto sob o ponto de vista matemático.

O índice de aninhamento mais conhecido na literatura que usa conceitos métricos é o índice temperatura, que estima o desvio de uma presença ou ausência não previstas em relação a uma isóclina, uma curva que separa duas regiões, de modo que a área abaixo da curva iguala-se à área sobre ela. Este índice tem muitos problemas que têm sido comentados na literatura [56, 57, 58]. Em nosso ponto de vista, os principais problemas com o índice temperatura é que a isóclina não tem uma equação analítica que a descreva e o índice não é apropriadamente normalizado.

O índice  $N_0$  é um estimador do número de lacunas na matriz de adjacência. O número de lacunas é definido como o número de ausências de espécies que não deveriam ocorrer em subconjuntos maiores. A matriz perfeitamente aninhada não tem lacunas, por definição.

O índice de discrepância de Brualdi e Sanderson foi introduzido pra superar o problema que surge quando duas colunas têm o mesmo número de células ocupadas, o que pode gerar ambiguidades. O índice de discrepância conta o número de elementos unitários que devem ser realocados para a matriz atingir o aninhamento perfeito.

Esses dois últimos índices compartilham um ponto em comum: eles estimam o quanto a matriz está longe do aninhamento perfeito. No entanto, eles não estimam o quanto um dado elemento discrepante está longe do padrão de aninhamento perfeito.

## 3.2 Temperatura Manhattan

Neste capítulo definimos a temperatura  $\tau$  da matriz e o índice de aninhamento  $\nu$  com o auxílio do conceito de métrica, uma função que define a distância entre elementos de um conjunto. A fim de eliminar a influência da assimetria entre os lados  $L_1$  e  $L_2$ , definimos as coordenadas cartesianas  $(x, y)$ :

$$x = i/L_1 \qquad y = j/L_2 \qquad (3.1)$$

Estas duas equações projetam os elementos  $a_{i,j}$  da matriz  $(L_1, L_2)$  no quadrado unitário. Esta transformação permite-nos preparar o problema para uma análise matemática.

Antes de aplicar o estimador de aninhamento aos dados de campo nós devemos empacotar a matriz. Esse processo define o canto de aninhamento e nós usamos o elemento  $a_{1,1}$  da matriz como o canto aninhado, como é usual na literatura desta área. Chamamos  $d_{x,y}$  a distância de um dado elemento a esse canto. O número de elementos "1" na matriz é  $N$

e usamos o número  $1 \leq k \leq N$  para indexar os elementos "1" na matriz de incidência. O objeto central na construção do aninhamento é a distância média da matriz de adjacência projetada no quadrado unitário:

$$d = \frac{1}{N} \sum_{k=1}^N d_{x_k, y_k} \quad (3.2)$$

Empregamos a distância de Manhattan  $d_{x_k, y_k} = x_k + y_k$  "porque ela é bastante prática para avaliar distâncias abstratas" em matrizes; ela é também utilizada, inclusive com vários outros nomes, na análise multivariada de composição de espécies em comunidades ecológicas e suas relações com vários fatores ambientais [59]. O intervalo da distância de Manhattan no quadrado unitário é  $0 \leq d_{x_k, y_k} \leq 2$ . Notamos que a distância de Manhattan é invariante perante permutações de linhas ou colunas com o mesmo número de elementos.

Esta propriedade da distância de Manhattan resolve um importante problema no conceito de aninhamento presente nos estimadores anteriores [5, 13]. Com efeito, quando a matriz é reajustada, algumas vezes a "posição de linhas ou colunas com a mesma soma marginal leva a diferentes valores de aninhamento". A métrica do aninhamento usando a distância de Manhattan evita tais ambiguidades. Vale lembrar que quando avaliamos a distância  $d$  estamos levando em consideração as informações de presenças e ausências inesperadas em relação ao aninhamento perfeito. Em ambos os casos  $d$  aumenta, presenças inesperadas são elementos com alto  $d_{x_k, y_k}$  e ausências inesperadas são elementos com baixo  $d_{x_k, y_k}$  que deixam de ser computados.

Para normalizar o índice  $\tau$  nós usamos duas matrizes artificiais: a matriz perfeitamente aninhada e a matriz randômica equiprovável. Estas duas matrizes têm o mesmo  $L_1$ ,  $L_2$  e  $\rho$  da matriz de adjacência, e elas funcionam como padrões de referência para definirmos apropriadamente os índices de aninhamento e temperatura. Em outras palavras, nós parametrizamos  $\tau$  com a ajuda das distâncias  $d_{nest}$  e  $d_{rand}$ . A primeira é a distância relacionada à matriz completamente aninhada e a segunda a distância média relacionada a um conjunto de matrizes aleatórias com distribuição uniforme, ambas projetadas no quadrado unitário. O índice de temperatura  $\tau$  foi inspirado no conceito de parâmetro de ordem. Este parâmetro, utilizado no estudo de transições de fase de sistemas termodinâmicos, pode ser definido como zero para um estado de desordem total e um para o estado de ordem completa. Como nossa analogia principal é com o conceito de temperatura, associado à desordem e não à ordem, o índice  $\tau$  é definido de modo inverso ao do parâmetro de ordem. Através de uma translação e uma normalização apropriadas da distância  $d$ , temos:

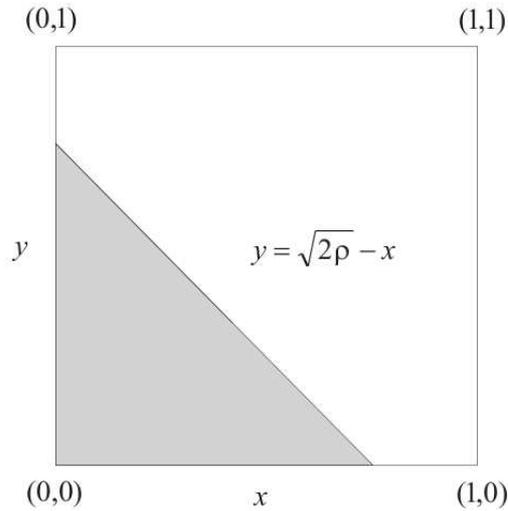


Figura 3.1: A representação de um padrão perfeitamente aninhado com área  $\rho$  no quadrado unitário. A linha oblíqua delimita o padrão aninhado. A equação Cartesiana desta linha é indicada na figura.

$$\tau = \frac{d - d_{nest}}{d_{rand} - d_{nest}} \quad (3.3)$$

Uma temperatura  $\tau$  igual a zero corresponde a um estado de mínima desordem, onde todos os elementos estão perfeitamente aninhados. No caso de  $\tau = 1$ , pelo contrário, temos um estado onde elementos não nulos da matriz estão distribuídos aleatoriamente com a mesma probabilidade. Muito embora esta matriz aleatória seja conhecida como modelo nulo  $R0$  na literatura de ecologia [6], ela não é usada aqui como um modelo nulo para avaliações baseadas em reamostragem, mas apenas como um padrão de referência para definir o parâmetro  $\tau$ . Nós mostramos com alguns detalhes como definir uma matriz perfeitamente aninhada e uma matriz aleatória, e procederemos ao cálculo analítico de  $d_{nest}$  e  $d_{rand}$ .

A matriz perfeitamente aninhada é tal que apresenta o mesmo  $L_1$ ,  $L_2$  e  $\rho$  da matriz original mas tem seus elementos maximamente aninhados. Para esta matriz especial a

distância  $d$  é menor ou igual à de qualquer outra matriz com o mesmo tamanho e ocupação. Determinamos  $d_{nest}$  analiticamente com o auxílio da integral das distâncias sobre a área do padrão aninhado. A figura 3.1 mostra a estrutura do padrão aninhado com área  $\rho$  quando projetado no quadrado unitário. O comprimento do lado do triângulo indicado na figura é  $\sqrt{2\rho}$ . A linha que delimita o triângulo é  $y = \sqrt{2\rho} - x$ . Para uma ocupação abaixo de  $\rho = 1/2$ , A distância média é obtida através da integral:

$$d_{nest} = \int_0^{\sqrt{2\rho}} \int_0^{\sqrt{2\rho}-x} d_{x,y} dy dx = \int_0^{\sqrt{2\rho}} \int_0^{\sqrt{2\rho}-x} (x+y) dy dx \quad (3.4)$$

o que resulta

$$d_{nest} = \frac{2}{3} \rho \sqrt{2\rho} \quad (3.5)$$

Uma integral dupla para  $1/2 < \rho < 1$  produz a relação

$$d_{nest} = 2\rho - 1 + \frac{2}{3} (1 - \rho) \sqrt{2(1 - \rho)} \quad (3.6)$$

A matriz aleatória que usamos na definição de  $\tau$  segue uma distribuição uniforme. Para calcular a distância média introduzimos a função densidade que segue a relação:

$$\begin{aligned} d_{rand} &= \int_0^1 \int_0^1 d_{x,y} \rho dx dy \\ d_{rand} &= \rho \int_0^1 \int_0^1 (x+y) dx dy = \rho \end{aligned} \quad (3.7)$$

Com o auxílio das equações (3.5) ou (3.6) e (3.7) encontramos a temperatura e o aninhamento para qualquer matriz com uma dada ocupação. As curvas de  $d_{rand}$  e  $d_{nest}$  podem ser vistas na figura 3.2.

### 3.3 Ocupação da Matriz e Otimização da Métrica

Segundo Ulrich e colaboradores [60], a métrica de aninhamento e sua performance estatística não devem, idealmente, ser afetadas pelo tamanho e pelo formato da matriz. A forma usual de eliminar o efeito destas duas propriedades da matriz no cálculo do aninhamento é a confrontação estatística dos resultados obtidos com modelos nulos. Nossa abordagem é bastante diferente. Nesta seção utilizamos a distância de Minkowski, que é uma generalização das métricas de Manhattan e euclideana, para eliminar a influência que diferentes graus de ocupação entre as matrizes exerceriam no nível de aninhamento. Isto garante que matrizes igualmente bem aninhadas mas com diferentes graus de ocupação

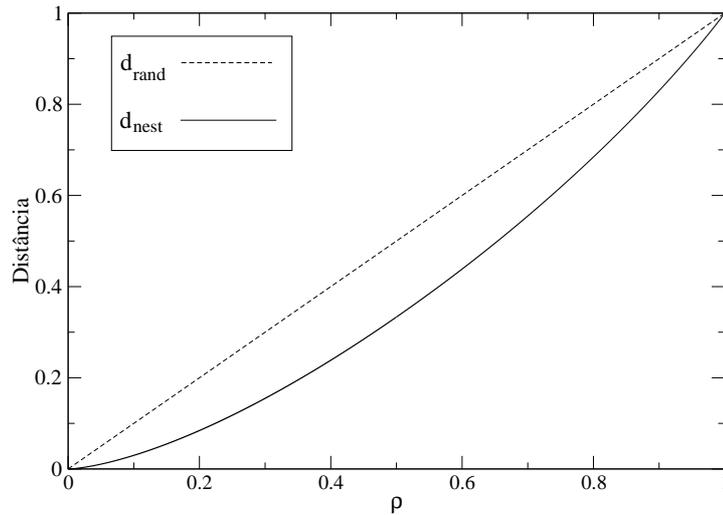


Figura 3.2: O comportamento de  $d_{nest}$  e  $d_{rand}$  versus  $\rho$ . As curvas obedecem à regra  $d_{nest} < d_{rand}$ , como esperado. Por construção, essas curvas não dependem do tamanho da matriz.

tenham o mesmo índice de aninhamento. As figuras 3.3 (a) e (b) mostram duas situações onde as matrizes estão completamente aninhadas mas têm valores diferentes de ocupação  $\rho$ : (a)  $\rho = 9/25$  e (b)  $\rho = 15/25$ . No caso da distância de Manhattan o conjunto de sítios com a mesma distância, chamados pontos isométricos, são linhas diagonais. Em outras palavras, os pontos isométricos são o conjunto de elementos  $a_{i,j}$  tais que  $c = i + j$  para uma dada constante  $c$ . É fácil verificar que estes pontos estão em uma diagonal da matriz.

Nós utilizamos um método sutil para eliminar a influência das diferentes ocupações no aninhamento. Provocamos uma distorção na linha isométrica para que ela "delimite a fronteira da ocupação  $\rho$ ". Na figura 3.4 nós representamos as linhas isométricas  $1 = (x^\chi + y^\chi)^{1/\chi}$  para  $\chi = 0.25, 0.5, 0.75, 1, 1.5, 2, 4, \text{ e } 10$ , como indicado na figura. Aqui nós usamos as coordenadas cartesianas  $(x, y)$  em vez das coordenadas  $(i, j)$  porque é

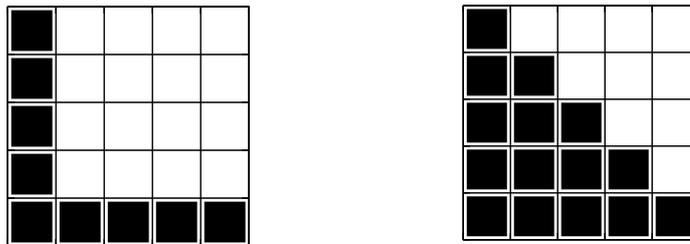


Figura 3.3: Estas duas matrizes estão igualmente bem aninhadas. Escolhemos um parâmetro métrico  $\chi$  adequado para produzir índices "similares" de aninhamento.

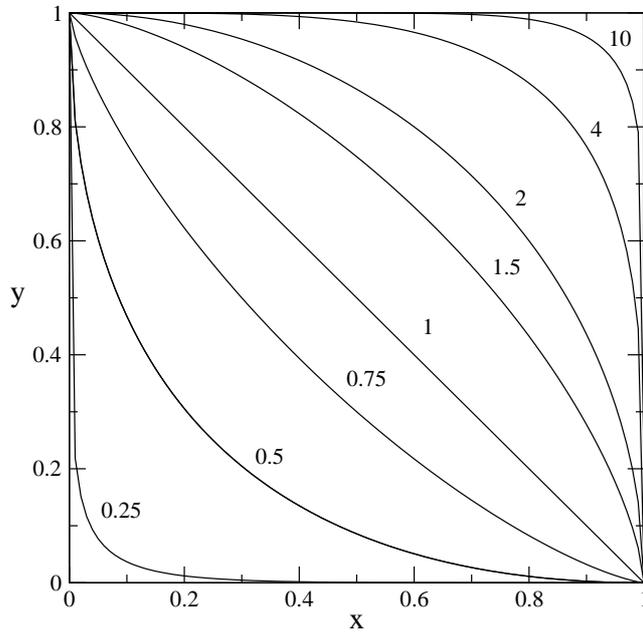


Figura 3.4: As curvas  $1 = (x^\chi + y^\chi)^{1/\chi}$  para vários valores de  $\chi$  indicados na figura. Escolhemos o melhor  $\chi$  de modo que delimite o padrão completamente aninhado para uma dada ocupação. Esta figura deve ser comparada com a figura 3.3. Nossa estratégia é encontrar um  $\chi$  ótimo de modo que a curva delimite o padrão aninhado.

mais conveniente trabalharmos analiticamente em um espaço cartesiano contínuo que no espaço discreto dos elementos da matriz. A escolha da região inferior e à esquerda da matriz para o aninhamento é feita para que o ponto cartesiano  $(1, 1)$  seja o foco do aninhamento, a fim de facilitar a análise. Podemos assim comparar a figura 3.3 com a figura 3.4. Determinamos a melhor curva para o padrão de aninhamento máximo por imposição de que a área sob a curva  $1 = (x^\chi + y^\chi)^{1/\chi}$  seja  $\rho$ . A equação implícita da linha isométrica pode ser escrita "explicitamente" como

$$y(x) = (1 - x^\chi)^{1/\chi}. \quad (3.8)$$

Deste modo a melhor linha isométrica que ajusta-se à ocupação  $\rho$  deve obedecer à relação de área:

$$\rho = \int_0^1 (1 - x^\chi)^{1/\chi} dx \quad (3.9)$$

A solução analítica desta equação encontra-se no Apêndice A e produz uma forma

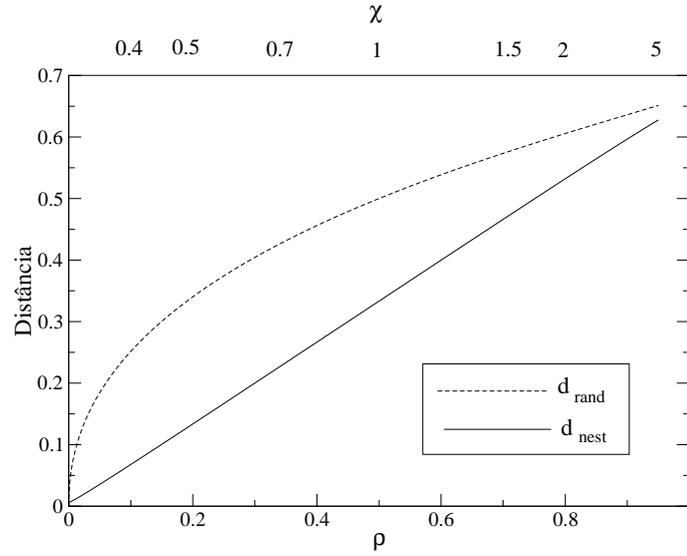


Figura 3.5: O comportamento de  $d_{nest}$  e  $d_{rand}$  versus  $\rho$ . As curvas obedecem a regra  $d_{nest} < d_{rand}$  como esperado. Alguns valores de  $\chi$  correspondentes a  $\rho$  estão indicados no topo da figura.

explícita para a relação entre  $\rho$  e  $\chi$ :

$$\rho = \frac{\left[ \Gamma\left(\frac{1}{\chi} + 1\right) \right]^2}{\Gamma\left(\frac{2}{\chi} + 1\right)}, \quad (3.10)$$

onde  $\Gamma$  é a função Gama. Esta equação determina a métrica, o parâmetro  $\chi$  da distância de Minkowski, através do valor experimental da ocupação  $\rho$ , o que significa que ela encontra o melhor valor da métrica para uma situação específica. É fácil observar na equação (3.10) que, como esperado, o caso  $\rho = 1/2$  corresponde a  $\chi = 1$ .

A distância da matriz aninhada é encontrada utilizando-se um procedimento similar ao descrito na seção anterior. Integramos a distância de Minkowski na região sob a linha isométrica dada pela equação (3.8):

$$d_{nest} = \int_0^1 \int_0^{(1-y^x)^{1/x}} d_{x,y} dx dy = \int_0^1 \int_0^{(1-y^x)^{1/x}} (x^x + y^x)^{1/x} dx dy \quad (3.11)$$

A distância da matriz aleatória é também encontrada de modo similar ao da seção anterior. A integral aqui é expressa como:

$$d_{rand} = \int_0^1 \int_0^1 \rho dx dy = \rho \int_0^1 \int_0^1 (x^x + y^x)^{1/x} dx dy \quad (3.12)$$

Ambas as integrais duplas acima foram resolvidas numericamente, mas podem ser

avaliadas também de modo analítico, o que foi feito no Apêndice A. A equação (3.11) foi resolvida da seguinte forma. A integração em relação à variável  $x$  produz uma função hipergeométrica com argumento  $(1 - y^x)^{1/\chi}$ . Esta função é expandida em uma série de potências e os termos são então integrados em relação à variável  $y$ . A segunda integração produz um novo conjunto de funções hipergeométricas. A equação (3.12) foi resolvida usando-se uma expansão similar, mas neste caso a segunda integração é trivial em virtude da simplicidade de seus limites. Na figura 3.5 nós mostramos ambas as curvas para  $d_{nest}$  e  $d_{rand}$  como função de  $\rho$ . Os valores de  $\chi$  correspondentes aos de  $\rho$  de acordo com (3.10) são mostrados na parte superior da figura. Podemos observar que para o caso  $\rho = 0.5$  este modelo degenera-se no caso da distância de Manhattan estudado na seção anterior. Por este motivo, para este caso especial, os valores de  $d_{nest} = 0.333$  e  $d_{rand} = 0.5$  são os mesmos nas figuras 3.3 e 3.5. Com os valores de  $d_{nest}$  e  $d_{rand}$  calculados deste modo obtemos os dois padrões de referência da equação (3.11) para a métrica otimizada e encontramos assim a temperatura de Minkowski.

### 3.4 Aplicação do novo índice de aninhamento a dados de campo

Para que a metodologia descrita acima seja aplicada no estudo de dados de campo, o programa computacional que desenvolvemos foi o MetricNest, escrito na linguagem do software livre Scilab e que pode ser facilmente convertido para o ambiente Matlab. O código é comentado no Apêndice ???. A primeira etapa dos cálculos consiste em empacotar a matriz, o que é feito através da ordenação das linhas e colunas, alternadamente e acordo com seu número de elementos, de modo que ao final deste procedimento tanto linhas mais abaixo e colunas mais à esquerda da matriz contenham sempre um número maior de elementos que as linhas mais acima e colunas mais à direita. Em seguida são calculadas as distâncias de Manhattan de cada um dos elemento da matriz empacotada. Estas distâncias são somadas para encontrarmos a distância  $d$  da matriz empacotada. Em seguida calculamos os padrões de referência  $d_{nest}$  e  $d_{rand}$ , as distâncias de Manhattan da matriz perfeitamente aninhada e da matriz randômica com a mesma ocupação  $\rho$  da matriz original. A partir destes valores calculamos a temperatura Manhattan  $\tau$ . O próximo passo é calcular a distância de Manhattan com os elementos projetados no quadrado unitário, o que é feito dividindo-se os dois fatores das distâncias de Manhattan pelas dimensões da matriz. Deste modo podemos encontrar o valor da distância de Manhattan projetada no quadrado unitário,  $\tau_p$ .

Em seguida o programa utiliza o segundo método que desenvolvemos e calcula a tem-

peratura de Minkowski da matriz. A metodologia básica é a mesma do caso Manhattan, mas aqui é encontrada a métrica de Minkowski  $\chi$  apropriada à ocupação  $\rho$  da matriz original. Em seguida são somadas as distâncias de Minkowski de todos os elementos da matriz para encontrarmos a distância  $d$ . São então realizadas as integrações numéricas apropriadas para encontrarmos  $d_{nest}$  e  $d_{rand}$ . Finalmente é então calculada a temperatura de Minkowski, chamada no código do programa de temperatura Isóclina. Todos estes passos são feitos tanto para a matriz empacotada original como para a sua projeção no quadrado unitário.

Os outros elementos do programa executam procedimentos de leitura e impressão de dados. O MetricNest tem a vantagem de ler diretamente as planilhas em formato .xls produzidas em campo e retornar os resultados dos cálculos em uma planilha com o mesmo formato.

## *Considerações Finais*

O aninhamento de uma matriz é a medida do quanto seus elementos podem ser empacotados sem lacunas. Nós propomos que para estimar o desvio de um padrão aninhado devemos considerar conceitos métricos em vez de utilizar estimativas baseadas em análise combinatória. Este trabalho abre uma nova perspectiva no estudo do aninhamento por apresentar um índice inspirado na termodinâmica e na métrica, que é fundamentado com o auxílio de dois padrões de referência bem postos: o perfeitamente aninhado e o uniformemente distribuído. Estes dois padrões possuem soluções analíticas. Nosso novo estimador de aninhamento  $\nu$  é baseado na análise métrica direta das distâncias sobre a matriz de incidência. Esperamos que este estimador venha a ser útil na discussão metodológica envolvendo aninhamento em matrizes ecológicas.

Nós reforçamos a idéia de que um bom índice de aninhamento deve quantificar o quanto uma dada matriz desvia-se do padrão de aninhamento perfeito. De fato, esta idéia está presente em muitos estudos tradicionais na área [2, 12, 13]. Mais recentemente, outras propriedades estão sendo atribuídas ao conceito de aninhamento e à sua medida [5]. No entanto, a visão tradicional de que uma medida de aninhamento deve expressar desvios do padrão de aninhamento completo preenche os requisitos de elegância, simplicidade e clareza.

Neste trabalho apresentamos uma exploração analítica do aninhamento em matrizes de presença e ausência que podem ser interpretadas como redes bipartidas. Adicionalmente, uma metodologia similar pode ser estendida a grafos. De um ponto de vista matemático, a diferença é que para grafos nós temos  $L_1 = L_2$  e a matriz de adjacência é simétrica. Acreditamos que esta técnica possa ser aplicada no contexto de biologia celular, onde espécies são substituídas por metabólitos e interações são interpretadas como coparticipação em uma reação bioquímica. Nesta área o coeficiente de aninhamento pode ser útil para quantificar a robustez de uma célula perante mudanças em sua "composição" metabólica produzida por distúrbios farmacológicos ou genômicos [61].

Em nossa visão, o aninhamento deve ser tratado como uma métrica e não como um conceito. Acreditamos que tão logo o aninhamento se torne um objeto matemático bem definido, maiores desenvolvimentos teóricos podem ser esperados nesta área. Questões como a relação entre aninhamento e abundância de espécies ou entre aninhamento e

riqueza de espécies em sítios podem então ser também postas em uma estrutura matemática mais consistente. No presente, estas questões são consideradas apenas através de simulações Monte Carlo. Uma vez que o aninhamento é um conceito abstrato envolvendo manipulações algébricas de matrizes, há um desequilíbrio nesta área. Existem poucos resultados analíticos comparados com estimativas computacionais e estatísticas. Esperamos que este trabalho ajude a superar a falta de estruturação matemática nesta área, o que é de grande interesse por parte da pesquisa corrente em biogeografia [60], biologia de conservação [7] e ecologia de comunidades [62, 4].

Como perspectivas para trabalhos futuros podemos:

- Publicar e Disponibilizar o programa MetricNest, a ferramenta computacional que desenvolvemos para o cálculo da distância de Manhattan, a fim de popularizá-la entre a comunidade de ecólogos e biólogos. Este programa apresenta vantagens em relação a outras ferramentas computacionais utilizadas para o cálculo do índice de aninhamento. Ele capta os dados diretamente da planilha de campo do pesquisador e retorna os resultados tanto da temperatura de Manhattan como da temperatura de Minkowski.

- Analisar a influência que a abundância de espécies exerce no aninhamento da rede como um todo, tanto no caso de interações entre espécies como ocorrência de espécies em sítios fragmentados, o que deve esclarecer o peso qualitativo e o quantitativo nestes tipos de relações ecológicas. Este trabalho está em andamento e resultados preliminares podem ser visto em nossos trabalhos [63, ?, ?].

- Realizar a efetiva aplicação de nosso estimador a dados disponíveis na literatura ou coletados por colaboradores, a fim de estabelecer em casos empíricos e concretos a eficiência de nosso método, e de contribuir para o melhor entendimento dos processos modelados por redes bipartidas e matrizes de incidência.

## *APÊNDICE A – Código do Programa MetricNest*

```

function manhattan(A); //Função Manhattan: calcula os parâmetros
                        //Manhattan e Manhattan com projeção.

A = nestinv(A); //Função nestinv: função que faz o aninhamento
                // da matriz no canto do elemento (1,1).

[x,y] = size(A); //Função size: função do Scilab que determina
                //as dimensões L1 e L2 da matriz.

d = 0;
N = 0;
//Cálculo Manhattan sem projeção;
for i= 1:x;
    for j = 1:y;
        if(A(i,j)== 1) then
            N = N + 1;
            d0 = i+j;
            d = d + d0;
        end
    end
end

//Cálculo das distâncias de Manhattan sem
// projeção de todos os elementos da matriz.

Rho = N/(x*y);
drand = (N*(x+y))/2;
Beta = (-1+sqrt(1+8*N))/2;
dnest = ((2/3)*Beta)+((Beta)^2)+(((Beta)^3)/3);
Tau = (d - dnest)/(drand - dnest);

//Cálculo dos parâmetros

printf("\n");

```

```

printf(" L1 = %i , L2 = %i",x,y);
printf("\n");
printf("\n___Manhattan sem projeção___");
printf("\nTau:%3f , dnest : %f , drand: %f, d: %f, N:%i",Tau,dnest,drand,d,N);
        //Imprime na tela os resultados.

//Cálculo Manhattan com projeção:
dp = 0;
for i = 1:x;
    for j = 1:y;
        if(A(i,j) == 1) then
            d1 = i/x + j/y; //Diferença no cálculo projetado: a distância é
                //dividida pelas dimensões L1 e L2 da matriz.
            dp = dp + d1;
        end
    end
end
        //Cálculo das distâncias de Manhattan com projeção
        //de todos os elementos da matriz.
dnestp = (2/3)*Rho*sqrt((2*Rho))*N*2;
        //linha modificada multiplicar por 2N
drandp = N;//linha modificada: N ao invés de Rho
Taup = (dp-dnestp)/(drandp-dnestp);
printf("\n");
printf("\n___Manhattan com projecao___");
printf("\nTau:%3f,dnest:%f,drand:%f,d:%f,N:%i",Taup,dnestp,drandp,dp,N);
P = [Tau dnest drand d N;Taup dnestp drandp dp N];
path = getcwd();
fprintfMat(path+'/manhattan.txt',P,'%5.5f');
        //Essa instrução serve apenas para gravar os dados
        //calculados num arquivo .txt no mesmo diretório
        //onde está o arquivo .xls da matriz.

endfunction

function [B] = nestinv(A);
        //Função nestinv que faz o aninhamento da Matriz.

```

```
[i,j]=size(A);
r = 0;
for y = 1:j;
    for x = 1:i;
        if( A(x,y) == 1 ) then
            r = r+1;
        end
    end
    //Definição do numero de elementos 1 em cada coluna.
    C(y) = r; //Vetor C que armazena o ranking das colunas:
              //o número de sítios ocupados em cada coluna.

    r = 0;
end

for y = 1:j;
    for x = j:-1:2;
        if(C(x-1)<C(x)) then
            aux = A(:,(x-1));
            A(:,(x-1)) = A(:,x);
            A(:,x) = aux;
            aux2 = C(x-1);
            C(x-1)=C(x);
            C(x)=aux2;
        end
    end
end
//Ordenacao das colunas.

r = 0;
for x = 1:i;
    for y = 1:j;
        if( A(x,y) == 1 ) then
            r = r+1;
        end
    end
    L(x) = r;
    r = 0;
end
//Definição do numero de elementos 1 em cada linha
```

```

for x = 1:i;
    for y = i:-1:2;
        if(L(y-1)<L(y)) then
            aux2 = A((y-1),:);
            A((y-1),:) = A(y,:);
            A(y,:) = aux2;
            aux = L(y-1);
            L(y-1) = L(y);
            L(y) = aux;
        end
    end
end
//Ordenação das linhas
B = A;
endfunction

deff(' [rho]=RHO(qsi)', 'rho =
(gamma((1/qsi))*gamma((1/qsi)+1))/(qsi*gamma((2/qsi)+1))');
//Definição da função para calcular o parâmetro
//Rho em função de qsi.
function [qsi] = qsi_rho(Rhoi)
    qsi = 0.02;
    for i = 1:1000;
        Rho = RHO(qsi);
        if(abs((Rhoi-Rho)/Rhoi)<=1D-6) then
            break;
        else
            E = (Rhoi-Rho)/Rhoi;
            qsi = qsi+(Rhoi*E);
        end
    end
endfunction
//Método recursivo para encontrar um valor de
//qsi para um dado Rho.

function [I] = int_simp(a,b,n,y,qui,func)
if modulo(n,2)<>0 then
    error("O valor de n deve ser par!")
end
h = (b - a)/n;

```

```
soma = 0
x0 = a
for i = 1:(n/2)
    xi = x0 + h;
    xii = xi + h;
    soma = soma + (h/3)*(func(x0,y,qui)+4*func(xi,y,qui)+func(xii,y,qui));
    x0 = xii
end
I = soma;

endfunction          //Método numérico (Simpson) para integração numérica.

function [z] = f(x,y,qui);

    z = (x^qui+y^qui)^(1/qui);

endfunction          //Função da isóclina.

function [z] = w(x,y,qui);

    [L1,L2] = size(A);
    z = ((x/L1)^qui+(y/L2)^qui)^(1/qui);

endfunction          //Função da isóclina projetada no quadrado unitário.

function [dnes] = dnest(qui,func)
    nx = 500;
    dx = 1/nx;
    dy = dx;
    i = 0;
    for y = 0:dy:1;
        i = i+1;
        a = 0;
        b = (1-y^qui)^(1/qui);
        Ix(i) = int_simp(a,b,nx,y,qui,func);
    end

    I = 0;
```

```
    for i = 1:nx;
        I = I + (dx/2)*(Ix(i)+Ix(i+1));
    end
    dnes = I;

endfunction          //Resolução da integral para calcular o valor de
                    //dnest da isóclina.

function [dran] = drand(qui,func)
rho = RHO(qui);
nx = 40;
dx = 1/nx;
dy = dx;
i = 0;
for y = 0:dy:1;
    i = i+1;
    a = 0;
    b = 1;
    Ix(i) = int_simp(a,b,nx,y,qui,func);
end

I = 0;
for i = 1:nx;
    I = I + (dx/2)*(Ix(i)+Ix(i+1));
end
dran = rho*I;

endfunction          //Resolução da integral para calcular o valor de
                    //drand da isóclina

function isocline(A)    //Semelhante ao caso de Manhattan.
A = nestinv(A);
[x,y] = size(A);
N = 0;
for i= 1:x;
    for j = 1:y;
        if(A(i,j)== 1) then
```

```

        N = N + 1;
    end
end
end
//Calcula número de sítios ocupados.
Rho = N/(x*y);
qsi = qsi_rho(Rho); //Calcula a densidade de ocupação RHO e a partir
//dela determina o qsi.

//Cálculo da distancia d sem projeção:
d = 0;
for i = 1:x;
    for j = 1:y;
        if(A(i,j) == 1) then
            d = d + (((i^qsi) + (j^qsi))^(1/qsi));
        end
    end
end
//Cálculo da distância Isoclina, parece muito com
//Manhattan. Deferente apenas no
//cálculo da distância d.

L = (x+y)/2; //Média aritmética das dimensões da matriz.

dnest_i = 2*L*N*dnest(qsi,f); //Linha modificada multiplicada por 2*L*N.
drand_i = 2*L*N*drand(qsi,f); //Linha modificada multiplicada por N.

Tau_i = (d - dnest_i)/(drand_i - dnest_i);
printf("\n");
printf("\n___ Isocline sem projeção___");
printf("\n");
printf("Tau:%3f,dnest:%f,drand:%f,d:%f,qsi:%f",Tau_i,dnest_i,drand_i,d,qsi);
printf("\n");

//calcula da distancia d com projeção
dp = 0;
for i = 1:x;
    for j = 1:y;
        if(A(i,j) == 1) then
            dp = dp + (((i/x)^qsi) + ((j/y)^qsi))^(1/qsi));
        end
    end
end

```

```

//Distância d Isóclina projetada.
    end
    end
end
L = (x+y)/2;          // linha acrescentada 28/04/2009 média aritmética das
                    //dimensões da matriz
dnest_ip = (dnest(qsi,w))*N*2*L;//linha modificada multiplicada por 2*L*N
drand_ip = (drand(qsi,w))*N*2*L;//linha modificada multiplicada por 2*L*N
                    //Essas funções parecem as da isóclina não projetada,
                    //mas é que multiplicada por 2NL e é chamada função w.

Tau_ip = (dp - dnest_ip)/(drand_ip - dnest_ip);
printf("\n___Isocline com projeção___");
printf("\n");
printf("\nTau:%3f,dnest:%f,drand:%f,d:%f,qsi:%f",Tau_ip,dnest_ip,drand_ip,dp,qsi);
P = [Tau_i dnest_i drand_i d qsi;Tau_ip dnest_ip drand_ip dp qsi];
path = getcwd();
fprintfMat(path+'/isocline.txt',P,'%5.5f');
                    //Salva os dados da isóclina num txt
endfunction

function metricnest (A)
    path = xgetfile(title='Escolha o arquivo .xls da matriz');
    sheet = readxls(path);
    A = sheet(1);
    A = A.value;          //Aqui pega a matriz no XLS e armazena na memória A
    printf("\nArquivo lido.")
    printf("\n")
    printf(path)
    printf('\n')
    manhattan(A);
    isocline(A);
endfunction          //Aqui ocorre a mágica, a função metricnest chama
                    //todas as outras.

```

## *Referências*

- [1] A. M. Almeida, Comunicação pessoal (2009).
- [2] P. J. Darlington, *Zoogeography: the geographical distribution of animals* (Wiley, New York, 1957).
- [3] T. M. Lewinsohn, P. I. Prado, P. Jordano, J. Bascompte, e J. M. Olesen, *Oikos* **113**, 174 (2006).
- [4] B. J. e J. P., *Annual Review of Ecology, Evolution, and Systematics* **38**, 567 (2007).
- [5] M. Almeida-Neto, P. Guimarães, P. R. Guimarães Jr, R. D. Loyola, e W. Ulrich, *Oikos* **117**, 1227 (2008).
- [6] W. Ulrich e N. J. Gotelli, *Ecology* **88**, 1824 (2007).
- [7] A. Cutler, *Conservation Biology* **5**, 496 (1991).
- [8] P. B. D. M. G. M. Wright, D. H., A. Cutler, e W. Atmar, *Oecologia* **113**, 1 (1998).
- [9] E. Burgos, H. Ceva, R. P. J. Perazzo, M. Devoto, D. Medan, M. Zimmermann, e A. M. Delbue, *Journal of Theoretical Biology* **249**, 307 (2007).
- [10] D. Medan, R. P. J. Perazzo, M. Devoto, E. Burgos, M. Zimmermann, H. Ceva, e A. M. Delbue, *Journal of Theoretical Biology* **246**, 510 (2007).
- [11] A. W. e P. B. D., *Oecologia* **246**, 373 (2007).
- [12] P. B. D. e A. W., *Biological Journal of the Linnean Society* **28**, 65 (1986).
- [13] R. A. Brualdi e J. G. Sanderson, *Oecologia* **119**, 256 (1999).
- [14] J. H. Lawton, *Ecology Letters* **6**, 69 (1979).
- [15] N. Martinez, B. Hawkins, H. Dawah, e B. Feifarek, *Ecology* **80**, 1044 (1999).
- [16] J. E. Cohen, *Food webs and niche space* (Princeton University Press, Princeton, 1978).
- [17] N. D. Martinez, *Ecological Monographs* **61** (1991).
- [18] P. Yodzis e K. Winemiller, *Oikos* **87**, 327 (1999).
- [19] N. D. Martinez, *The American Naturalist* **144**, 935 (1994).
- [20] K. Winemiller, E. Pianka, L. Vitt, e A. Joern, *The American Naturalist* **158**, 193 (2001).
- [21] J. Memmott, G. H.C.J., e I. Gauld, *Journal of Animal Ecology* **63**, 521 (1994).

- [22] C. Müller, I. Adriaanse, R. Belshaw, e H. C. J. Godfray, *Journal of Animal Ecology* **68**, 346 (1999).
- [23] G. R. Valladares e A. Salvo, *Environmental Entomology* **28**, 539 (1999).
- [24] A. S. Rott e H. C. J. Godfray, *Journal of Animal Ecology* **69**, 274 (2000).
- [25] K. Schönrogge e M. J. Crawley, *Journal of Animal Ecology* **69**, 841 (2000).
- [26] D. Wright, B. Patterson, G. Mikkelsen, A. Cutler, e W. Atmar, *Oecologia* **113**, 1 (1998).
- [27] J. Bascompte, P. Jordano, C. J. Melián, e J. M. Olesen, *Proceedings of the National Academy of Sciences USA* **100**, 9383 (2003).
- [28] J. Dunne, R. Williams, e N. Martinez, *Ecology Letters* **5**, 558 (2002).
- [29] C. Melián e J. Bascompte, *Ecology Letters* **5**, 705 (2002).
- [30] R. J. Putman, *Community Ecology* (Chapman and Hall, London, 1994).
- [31] L. Dicks, S. Corbet, e R. Pywell, *Journal of Animal Ecology* **71**, 32 (2002).
- [32] S. Maslov e K. Sneppen, *Science* **296**, 910 (2002).
- [33] A. Barabási e R. Albert, *Science* **286**, 509 (1999).
- [34] A. Barabási e E. Bonabeau, *Scientific American Brasil* **13**, 64 (2003).
- [35] R. Albert, H. Jeong, e A. Barabási, *Nature* , 378 (406).
- [36] J. Montoya e R. Solé, *Oikos* **102**, 614 (2003).
- [37] D. Raffaelli e S. Hall, *Journal of Animal Ecology* **61**, 551 (1992).
- [38] P. Jordano, J. Bascompte, e J. Olesen, *Ecology Letters* **6**, 69 (1979).
- [39] E. A. Bernays e R. Chapman, *Host-plant selection by phytophagous insects* (Chapman and Hall, New York, 1994).
- [40] K. Schönrogge e M. J. Crawley, *The coevolutionary process* (University of Chicago Press, Chicago, 1994).
- [41] K. J. Gaston, D. Reavey, e G. R. Valladares, *Ecological Entomology* **17**, 86 (1992).
- [42] C. Fonseca e G. Ganade, *Ecological Monographs* **49**, 33 (1979).
- [43] S. L. Pimm e J. H. Lawton, *Journal of Animal Ecology* **49**, 879 (1980).
- [44] C. Fonseca e G. Ganade, *Journal of Animal Ecology* **65**, 339 (1996).
- [45] A. Salvo, *Diversidad y estructura en comunidades de parasitoides (Hymenoptera) de minadores de hojas (Diptera: Agromyzidae)*, Tese de doutorado, Facultad de Ciencias Exactas Físicas y Naturales. Universidad Nacional de Córdoba, Argentina, 1996.

- [46] B. Mccune e M. J. Mefford, *Multivariate analysis of ecological data, Version 4* (MjM Software Design, Gleneden Beach, 1999).
- [47] P. Prado, *Determinantes de ampla escala das associações entre insetos e plantas: Tefritídeos (Diptera) endófagos de capítulos de Asteráceas*, Tese de doutorado, Instituto de Biologia, Unicamp, 1999.
- [48] A. Almeida, *Biogeografia de interações entre Eupatorieae (Asteraceae) e insetos endófagos de capítulos na Serra da Mantiqueira*, Tese de doutorado, Instituto de Biologia, Unicamp, 2001.
- [49] J. e. a. Martiny, *Nature Reviews Microbiology* **4**, 102 (2006).
- [50] R. H. MacArthur, *The Theory of Island Biogeography* (Princeton University Press, Princeton, 1967).
- [51] J. E. M. Watson, J. Fischer, J. C. Ingram, e R. J. Whittaker.
- [52] R. J. Whittaker e J. M. Fernández-Palacios, *Island Biogeography: Ecology, Evolution and Conservation* (Oxford University Press, Oxford, 2007).
- [53] G. Chartran, *Introductory Graph Theory* (Prindle Weber and Schmidt, Boston, 1985).
- [54] J. Newman, *Scientific American* **189**, 66 (1953).
- [55] E. F. Krause, *Taxicab Geometry: An Adventure in Non-Euclidean Geometry* (Dover, New York, 1986).
- [56] M. Almeida Neto, P. R. Guimarães Jr, e T. M. Lewinsohn, *Oikos* **116**, 716 (2007).
- [57] J. Fischer e D. B. Lindenmayer, *Oikos* **99**, 193 (2002).
- [58] J. E. Moore e R. K. Swihart, *Oecologia* **152**, 1432 (2007).
- [59] P. Legendre e L. Legendre, *Numerical Ecology* (Elsevier, Amsterdam, 1988).
- [60] W. Ulrich, M. Almeida Neto, e M. J. Gotelli, *Oikos* **118**, 3 (2009).
- [61] P. D'haeseleer, S. Liang, e R. Somogyi, *Bioinformatics* **16**, 707 (2001).
- [62] T. Lewinsohn, P. Prado, e A. Almeida, *Inventários bióticos centrados em recursos: insetos fitófagos e plantas hospedeiras em Conservação da Biodiversidade em Ecossistemas Tropicais ed. Garay, I. e Dias, B. F. S.* (Vozes, Petrópolis, 2001).
- [63] G. Corso, A. I. L. Araújo, e A. M. Almeida, *Condensed Matter* , arXiv:0803.0007 (2008).
- [64] A. I. L. Araújo, G. Corso, A. M. Almeida, e T. Lewinsohn, *Physica A* (2009).
- [65] G. Gruenhage, *Generalized Metric Spaces. Handbook of Set-Theoretic Topology* (North-Holland, Amsterdam, 1984).
- [66] B. D. Patterson, *Conservation Biology* **1**, 323 (1987).
- [67] K. Schönrogge e M. J. Crawley, *Oecologia* **32**, 85 (1978).

# Livros Grátis

( <http://www.livrosgratis.com.br> )

Milhares de Livros para Download:

[Baixar livros de Administração](#)

[Baixar livros de Agronomia](#)

[Baixar livros de Arquitetura](#)

[Baixar livros de Artes](#)

[Baixar livros de Astronomia](#)

[Baixar livros de Biologia Geral](#)

[Baixar livros de Ciência da Computação](#)

[Baixar livros de Ciência da Informação](#)

[Baixar livros de Ciência Política](#)

[Baixar livros de Ciências da Saúde](#)

[Baixar livros de Comunicação](#)

[Baixar livros do Conselho Nacional de Educação - CNE](#)

[Baixar livros de Defesa civil](#)

[Baixar livros de Direito](#)

[Baixar livros de Direitos humanos](#)

[Baixar livros de Economia](#)

[Baixar livros de Economia Doméstica](#)

[Baixar livros de Educação](#)

[Baixar livros de Educação - Trânsito](#)

[Baixar livros de Educação Física](#)

[Baixar livros de Engenharia Aeroespacial](#)

[Baixar livros de Farmácia](#)

[Baixar livros de Filosofia](#)

[Baixar livros de Física](#)

[Baixar livros de Geociências](#)

[Baixar livros de Geografia](#)

[Baixar livros de História](#)

[Baixar livros de Línguas](#)

[Baixar livros de Literatura](#)  
[Baixar livros de Literatura de Cordel](#)  
[Baixar livros de Literatura Infantil](#)  
[Baixar livros de Matemática](#)  
[Baixar livros de Medicina](#)  
[Baixar livros de Medicina Veterinária](#)  
[Baixar livros de Meio Ambiente](#)  
[Baixar livros de Meteorologia](#)  
[Baixar Monografias e TCC](#)  
[Baixar livros Multidisciplinar](#)  
[Baixar livros de Música](#)  
[Baixar livros de Psicologia](#)  
[Baixar livros de Química](#)  
[Baixar livros de Saúde Coletiva](#)  
[Baixar livros de Serviço Social](#)  
[Baixar livros de Sociologia](#)  
[Baixar livros de Teologia](#)  
[Baixar livros de Trabalho](#)  
[Baixar livros de Turismo](#)