

Matemática para o Século XXI

Volume III

Teoria das Redes



Em 1736, Leonhard Euler recebeu um problema que circulava entre os habitantes de Königsberg, uma cidade da Prússia que hoje se chama Kaliningrado e fica na Rússia.

A cidade era cortada pelo rio Pregel, que formava duas ilhas no seu interior. Sete pontes conectavam as ilhas às margens e entre si. A pergunta que os moradores faziam uns aos outros, como desafio, era esta: é possível fazer um passeio pela cidade cruzando cada uma das sete pontes exatamente uma vez?

Euler provou que não. Mas o que o tornou um dos momentos fundadores da matemática moderna não foi a resposta — foi a forma como ele chegou a ela. Euler percebeu que os detalhes geográficos eram irrelevantes. O que importava era apenas a estrutura das conexões: quais massas de terra estavam conectadas a quais pontes. Ele abstraiu o problema numa representação de pontos e linhas — o que hoje chamamos de grafo — e provou que um passeio desse tipo só é possível quando no máximo dois dos pontos têm um número ímpar de conexões.

Com esse argumento, Euler inventou a teoria dos grafos. E sem saber, inaugurou a matemática das redes — uma área que no século XXI se tornaria central para compreender desde a propagação de vírus até o funcionamento da internet, desde a organização do cérebro humano até a estrutura do poder político.

O que uma rede revela

Há duas formas de estudar um sistema complexo.

A primeira é estudar as suas partes. Analisar cada neurônio individualmente, cada proteína separadamente, cada transação financeira por si mesma. Essa abordagem produz conhecimento profundo sobre os componentes — mas frequentemente falha em explicar o comportamento do sistema como um todo.

A segunda é estudar as conexões. Não o que cada parte é, mas como as partes se relacionam — quem se conecta a quem, com que intensidade, em que direção, com que frequência. Essa abordagem sacrifica o detalhe dos componentes, mas revela padrões que a análise componente a componente nunca veria.

A teoria das redes é a matemática dessa segunda abordagem. Ela pergunta: o que a estrutura das conexões nos diz sobre o comportamento coletivo?

As respostas são, consistentemente, surpreendentes.

Uma rede com apenas 5% das conexões possíveis pode ser um mundo pequeno — onde qualquer nó está a poucos passos de qualquer outro. Uma rede onde alguns nós têm milhares de conexões e a maioria tem apenas uma ou duas exibe comportamentos radicalmente diferentes de uma rede onde as conexões são distribuídas uniformemente. A remoção de um único nó bem conectado pode fragmentar uma rede inteira. A adição de uma única conexão entre dois grupos isolados pode transformar a dinâmica de todo o sistema.

Essas propriedades emergem da estrutura, não dos componentes. E a matemática das redes é a linguagem para descrevê-las com precisão.

As métricas que revelam a estrutura

Antes de entender como uma rede se comporta, é preciso saber descrevê-la.

O grau de um nó — o número de conexões que ele tem — é a medida mais simples e uma das mais informativas. Em redes sociais, o grau mede quantas conexões diretas uma pessoa tem. Em redes de citações científicas, mede quantas vezes um artigo foi citado. Em redes de infraestrutura, mede quantos caminhos passam por um ponto.

Mas o grau sozinho não captura a posição estratégica de um nó na rede. Um nó com muitas conexões pode ser periférico se todas as suas conexões forem entre si. Um nó com poucas conexões pode ser central se for a única ponte entre dois grupos grandes.

As métricas de centralidade quantificam essa posição estratégica de formas diferentes. A centralidade de proximidade mede o quão eficientemente um nó se comunica com todos os outros. A centralidade de intermediação mede quantos caminhos entre outros pares de nós passam por ele — uma medida de controle sobre o fluxo de informação. A centralidade de autovetor — que inclui o PageRank que o Google usou para ordenar páginas da web — mede a influência de um nó em função da influência dos nós a que está conectado.

Juntas, essas métricas formam um vocabulário para descrever a posição de cada nó numa rede — quão conectado está, quão próximo está dos outros, quão essencial é para manter a rede conectada, quão influente é em função de com quem se conecta.

Como as redes reais são formadas

Durante décadas, o modelo padrão para redes era o de Erdős e Rényi: colocar n nós e conectar cada par com probabilidade p , independentemente. O resultado é uma rede aleatória onde as conexões são distribuídas de forma aproximadamente uniforme — a maioria dos nós tem um número de conexões próximo da média, com poucos nós muito mais ou muito menos conectados.

Esse modelo é matematicamente elegante e teoricamente bem compreendido. Mas as redes reais não se comportam assim.

Em 1998, Duncan Watts e Steven Strogatz publicaram um artigo que descrevia um fenômeno que a intuição já suspeitava mas a matemática ainda não havia formalizado: redes do mundo real tendem a ser simultaneamente localmente densas e globalmente compactas. Os seus amigos tendem a conhecer uns aos outros — alta densidade local — e ao mesmo tempo qualquer pessoa está a poucos graus de separação de qualquer outra — pequena distância global. É o que ficou conhecido como o mundo pequeno.

Um ano depois, Albert-László Barabási e Réka Albert descreveram outro padrão que os dados revelavam consistentemente: a maioria dos nós tem poucas conexões, mas alguns nós — os hubs — têm um número de conexões vastamente superior à média. A distribuição de graus segue uma lei de potência. A internet, as redes de citações científicas, as redes de colaboração em filmes, as redes metabólicas — todas exibiam essa estrutura.

E a razão, Barabási e Albert mostraram, é simples: as redes crescem, e os nós novos preferem se conectar a nós que já têm muitas conexões. A popularidade gera mais popularidade. A riqueza atrai mais riqueza. É a lógica da ligação preferencial — e produz, inevitavelmente, uma rede com poucos hubs muito conectados e muitos nós periféricos.

Esses dois modelos — pequeno mundo e escala livre — transformaram a nossa compreensão de como sistemas complexos reais são organizados. E este volume desenvolve ambos com rigor matemático completo, incluindo as suas propriedades e limitações.

Quando a estrutura determina o destino

A pergunta mais importante da teoria das redes não é descritiva — como a rede é — mas dinâmica: como a estrutura da rede determina o que acontece nela.

Uma epidemia se propaga de forma fundamentalmente diferente numa rede de mundo pequeno e numa rede aleatória. Numa rede com hubs — poucos nós com muitas conexões — o vírus que infectar um hub cedo se espalhará muito mais rápido do que numa rede uniforme. A estratégia de vacinação ótima numa rede com hubs é radicalmente diferente da estratégia ótima numa rede uniforme — e ignorar essa diferença produz políticas de saúde pública sistematicamente subótimas.

A informação se difunde de forma diferente da desinformação, mesmo numa rede idêntica, porque os mecanismos de adoção são diferentes. Uma notícia que se propaga independentemente — cada pessoa que a lê pode transmiti-la sem precisar que os seus amigos também a tenham lido — segue uma dinâmica de broadcast. Uma prática social que só é adotada quando uma fração suficiente dos seus contatos também a adotou — fumar, usar um aplicativo, participar de uma manifestação — segue uma dinâmica de contágio complexo, com limiares e cascatas.

As transições de fase em redes — os momentos em que uma pequena mudança num parâmetro produz uma mudança qualitativa dramática no comportamento do sistema — são análogos às transições de fase na física. A emergência do componente gigante numa rede aleatória à medida que a probabilidade de conexão aumenta é matematicamente idêntica à transição de percolação na física estatística.

E os pontos de não-retorno — os momentos em que um sistema cruza um limiar além do qual a dinâmica muda irreversivelmente — são previsíveis a partir da estrutura da rede, mesmo quando são impossíveis de prever a partir do comportamento local dos seus componentes.

O paradoxo da amizade

Antes de avançar, vale a pena parar num resultado que ilustra perfeitamente o tipo de insight que a teoria das redes oferece — simples de enunciar, surpreendente de verificar, e com implicações que a intuição nunca anteciparia.

Em 1991, o sociólogo Scott Feld provou matematicamente o seguinte: em qualquer rede, a média do número de amigos dos seus amigos é maior ou igual ao número médio de amigos de uma pessoa.

Em linguagem simples: os seus amigos têm, em média, mais amigos do que você.

Isso parece impossível. Se A é amigo de B, e B é amigo de A, a relação é simétrica — como A pode ter sistematicamente menos amigos que B?

A resposta está num viés de amostragem sutil: quando você enumera os amigos dos seus amigos, os nós com mais conexões aparecem mais vezes — proporcionalmente ao número das suas conexões. As pessoas muito conectadas têm mais probabilidade de ser amigos seus, e têm mais amigos do que a média. A amostra que você observa quando olha para os seus amigos é enviesada para os nós mais conectados.

O mesmo paradoxo aparece em praticamente todos os domínios onde redes estão presentes. Os aeroportos em que você faz conexões têm mais rotas do que o aeroporto médio. Os artigos que você lê têm mais citações do que o artigo médio. Os colegas de trabalho que você conhece têm mais conexões profissionais do que o profissional médio.

É um resultado de álgebra elementar com consequências filosóficas profundas: a experiência subjetiva de estar numa rede é sistematicamente diferente da realidade estatística da rede. O que vemos quando olhamos para as nossas conexões é uma amostra distorcida — e saber que está distorcida, e em que direção, muda a forma como interpretamos a nossa posição no mundo.

Redes e os volumes seguintes

A teoria das redes não é um tópico isolado dentro desta coleção. É uma estrutura conceitual que atravessa todos os volumes seguintes.

No Volume IV, os modelos baseados em agentes são frequentemente definidos sobre redes — os agentes interagem com vizinhos numa grade, num grafo aleatório, ou numa rede de mundo pequeno. A estrutura da rede determina quais agentes influenciam quais, e portanto determina profundamente a dinâmica emergente do sistema.

No Volume V, os sistemas dinâmicos em redes — como estados de opinião se propagam, como sincronização emerge em osciladores acoplados, como cascatas de falha se propagam em infraestruturas — são uma extensão natural da dinâmica em sistemas de baixa dimensão para sistemas com estrutura relacional explícita.

No Volume VI, as redes neurais artificiais são, literalmente, redes — grafos direcionados onde os nós são neurônios e as arestas são conexões com pesos. A teoria da propagação de sinal em redes, que este volume desenvolve no contexto de epidemias e difusão de informação, é estruturalmente análoga à teoria da propagação de ativação em redes neurais.

E a ideia de que a estrutura das conexões determina o comportamento coletivo — que é o tema central deste volume — é também o tema central do aprendizado de máquina: o que a rede aprendeu está codificado não nos neurônios individuais, mas no padrão das conexões entre eles.

Como usar este volume

Este volume pode ser lido de duas formas.

A primeira é linear — do conceito mais simples ao mais complexo, das métricas estáticas à dinâmica sobre redes, dos modelos teóricos às aplicações práticas. Essa é a forma recomendada para quem está encontrando a teoria das redes pela primeira vez.

A segunda é temática — começando pelo tópico mais relevante para o seu contexto imediato e usando os outros capítulos como referência. Um epidemiologista pode começar pelos processos de contágio. Um analista de dados pode começar pelas métricas de centralidade e pelas ferramentas computacionais. Um economista pode começar pelas redes de dependência e pelos pontos de não-retorno.

Ambas as formas funcionam porque cada capítulo foi escrito para ser inteligível de forma relativamente autônoma, com referências explícitas aos conceitos de capítulos anteriores quando necessários.

Os exercícios incluem análises de redes reais — a rede de colaboração científica, a rede de rotas aéreas, a rede de seguimento no Twitter — usando as bibliotecas NetworkX e Gephi. A teoria é importante. Mas há algo que só se percebe quando se coloca uma rede real numa ferramenta de visualização e se vê a estrutura emergir dos dados: a sensação de que se está vendo algo que estava lá o tempo todo, invisível, esperando a linguagem certa para ser descrito.

Uma última palavra

Em 1999, Mark Granovetter publicou um dos artigos mais citados das ciências sociais — *A Força dos Laços Fracos*. A sua tese, contra-intuitiva e matematicamente demonstrável, era que os laços mais importantes para a mobilidade profissional e a difusão de informação não são os laços fortes — os amigos

próximos, a família, os colegas de confiança — mas os laços fracos: os conhecidos ocasionais, os ex-colegas perdidos de vista, as conexões de segundo grau.

A razão é estrutural. Os seus laços fortes tendem a conectá-lo a pessoas que já conhecem as mesmas pessoas que você — formam um grupo denso e fechado. Os seus laços fracos, por outro lado, conectam-no a partes da rede social que de outra forma seriam inacessíveis. A informação que chega por um laço fraco é provável que seja nova. A informação que chega por um laço forte é provável que você já conheça.

Este resultado transformou a forma como sociólogos pensam sobre capital social. E é um exemplo perfeito do que a teoria das redes faz melhor: pegar uma intuição vaga — as conexões importam — e transformá-la numa afirmação precisa, matematicamente demonstrável, empiricamente verificável, e que revela algo sobre o mundo que nenhuma outra abordagem revelaria.

As redes estão em todo lugar. A matemática para descrevê-las existe há menos de três séculos. E ainda estamos aprendendo o que ela pode revelar.

Este volume é uma introdução a essa revelação.

Sumário

1. Introdução — A Matemática das Conexões

2. Poder Computacional

Muitas Mentes Trabalhando Juntas

Classes de Big-O — Níveis de Poder Computacional

Hierarquia das Complexidades

Computação Distribuída na Tomada de Decisão

Coordenação vs. Computação

Algoritmos Distribuídos

Complexidade

3. Elementos das Redes

Nós e *links*

Caminhos (*path*) e Ciclos

Matriz de Adjacência

Cliques

Definição formal de rede

Métricas de Rede

Hubs e *Clusters*

Grau (*degree*)

Centralidade

Centralidade de Grau — Os *Hub* de Conexão

Closeness centrality (Centralidade de Proximidade) — Os Comunicadores Eficientes

Betweenness centrality (centralidade de intermediação) — As Pontes

Centralidade de Autovetor (*Eigenvector*) — A Influência dos Influenciadores

PageRank

Fórmula do PageRank

Coeficiente de *Clustering* (*Clustering Coefficient*)

Coeficiente de *Clustering* Local

Coeficiente de *Clustering* Médio

Usando o Python

Percorrendo a Rede

Comprimento do Caminho (*Path Length*)

Densidade da Rede

Comunidade e Modularidade em Redes

O Algoritmo de Girvan–Newman

4. A Natureza Dual das Redes

Uma Estrutura, Dois Mundos: A Rede em Estrela

- Grau Médio:
- Comprimento Médio do Caminho (APL):
- *Betweenness centrality*
- Coeficiente de *clustering*:

Aplicações para Redes em Estrela

5. Visualização e Análise de Redes

Gephi

NetLogo

Python

6. Redes Aleatórias (*Random Graphs*)

Modelo Erdős–Rényi (ER) – A Rede Puramente Aleatória

Modelo $G(n, M)$

Modelo $G(n, p)$

Emergência de Componentes Gigantes

Rede Aleatória no NetLogo

Modelo Barabási–Albert (BA) – (“os ricos ficam mais ricos”)

Modelo Watts–Strogatz (WS) – O Pequeno Mundo

Rede SW no NetLogo.

Exemplo

7. Análise Prática de Redes

8. Dinâmica das Redes

Transições de Fase e *Tipping Points*

Ergodicidade

Não ergodicidade

Dependência Do Caminho (*Path Dependence*)

O Processo de Pólya (Urna de Pólya)

Da Urna às Redes — O Mesmo Mecanismo

Tomada de Decisão sob Incerteza

Simulações no Python

Propagação nas Redes

Error! No text of specified style in document.

Processo de *Broadcast* (transmissão)

Processo de Difusão

Processo de Contágio

9. O Paradoxo da Amizade

10. Passos para Analisar uma Rede

11. Exercícios sobre Redes

Exercício 1 — Análise de Redes Sociais Políticas (Twitter/X)

Exercício 2 — Centralidades em Redes de Transporte

10. Epílogo

Bibliografia

Error! No text of specified style in document.