Pendahuluan

Statistik sering dimulai dari tabel: baris sebagai unit pengamatan, kolom sebagai variabel, dan angka sebagai nilai yang diukur. Bentuk tabel ini sangat kuat, tetapi tidak semua informasi statistik hidup secara alami di dalam baris dan kolom yang saling berdiri sendiri. Banyak data modern justru berisi relasi: siapa berinteraksi dengan siapa, variabel mana bergantung pada variabel lain, wilayah mana bertetangga dengan wilayah lain, atau gen mana berhubungan dengan gen lain dalam suatu proses biologis. Ketika relasi menjadi objek utama, teori graf menyediakan bahasa matematis yang tepat.

Secara paling dasar, sebuah graf adalah cara formal untuk mencatat objek dan hubungan antarobjek. Objek disebut simpul atau vertex, sedangkan hubungan disebut sisi atau edge. Notasi yang akan kita pakai adalah

$$G = (V, E),$$

dengan V sebagai himpunan simpul yang tidak kosong dan E sebagai himpunan sisi. Notasi ini adalah notasi standar dalam teori graf modern; buku-buku dasar seperti Diestel dan West memakai kerangka yang sama untuk membangun teori graf dari objek diskret yang sederhana menuju struktur yang jauh lebih kaya (Diestel, 2017; West, 2001).

Sebagai contoh pertama, bayangkan lima responden survei:

$$V = \{A, B, C, D, E\}.$$

Jika kita ingin mencatat siapa mengenal siapa, maka sebuah sisi dapat menyatakan relasi “saling mengenal”. Misalnya,

$$E = \{\{A, B\}, \{A, C\}, \{B, D\}, \{D, E\}\}.$$

Graf ini mengatakan bahwa A mengenal B, A mengenal C, B mengenal D, dan D mengenal E. Di sini sisi ditulis sebagai pasangan tak berurutan, misalnya A,B\, karena relasi “saling mengenal” dianggap simetris: jika A mengenal B, maka B mengenal A.

Contoh ini terlihat sederhana, tetapi pola berpikirnya sangat umum. Dalam statistik jaringan, simpul dapat berupa individu, rumah tangga, perusahaan, artikel ilmiah, gen, akun media sosial, atau wilayah geografis. Sisi dapat berarti persahabatan, transaksi, kutipan, interaksi biologis, komunikasi, kedekatan spasial, atau korelasi yang melewati ambang tertentu. Newman membahas secara luas bagaimana jaringan seperti ini muncul dalam sistem sosial, teknologi, biologi, dan informasi (Newman, 2018), sedangkan Kolaczyk menekankan bahwa jaringan dapat diperlakukan sebagai objek data statistik yang memerlukan ringkasan, pemodelan, estimasi, dan validasi (Kolaczyk, 2009).

Namun buku ini tidak hanya membahas graf sebagai gambar. Gambar membantu intuisi, tetapi teori graf bekerja karena graf adalah objek matematis. Jika dua orang menggambar graf yang sama dengan posisi simpul berbeda di kertas, grafnya tetap sama selama simpul dan sisinya sama. Inilah perbedaan penting antara representasi visual dan struktur matematis. Dalam statistik, perbedaan ini sangat penting: kita tidak ingin kesimpulan berubah hanya karena tata letak gambar berubah.

Ada tiga alasan utama mengapa teori graf penting untuk statistik.

Pertama, graf memberi bahasa untuk struktur dependensi. Dalam banyak model statistik, kita tidak hanya peduli pada nilai variabel, tetapi juga pada hubungan antarvariabel. Misalnya, dalam model grafis probabilistik, simpul dapat merepresentasikan variabel acak, dan sisi dapat merepresentasikan bentuk dependensi atau ketergantungan bersyarat. Ide ini menjadi dasar bagi Bayesian network, Markov random field, dan banyak model probabilistik modern; Lauritzen memberikan perlakuan matematis klasik tentang bagaimana graf dipakai untuk menyatakan struktur dependensi probabilistik (Lauritzen, 1996).

Sebagai contoh, misalkan kita mempelajari tiga variabel: hujan, jalan basah, dan kemacetan. Variabel “hujan” dapat memengaruhi “jalan basah”, dan “jalan basah” dapat memengaruhi “kemacetan”. Struktur ini dapat digambarkan sebagai graf berarah:

$$\text{Hujan} \to \text{Jalan Basah} \to \text{Kemacetan}.$$

Panah di sini bukan sekadar dekorasi. Panah menyatakan arah hubungan yang dimodelkan. Dalam konteks kausalitas, graf berarah asiklik atau directed acyclic graph sering dipakai untuk menyatakan asumsi sebab-akibat, membedakan asosiasi dari klaim kausal, dan merumuskan kondisi identifikasi seperti back-door criterion; kerangka ini dibahas secara sistematis oleh Pearl (2009).

Kedua, graf memberi cara untuk memahami struktur data yang tidak independen. Banyak metode statistik klasik lebih mudah diterapkan ketika pengamatan diasumsikan independen. Tetapi dalam jaringan sosial, pengamatan antarindividu sering tidak independen: perilaku seseorang mungkin berkaitan dengan perilaku teman, kolega, atau tetangganya. Dalam data spasial, wilayah yang bertetangga sering memiliki karakteristik serupa. Dalam data biologis, ekspresi gen dapat terorganisasi dalam jaringan interaksi. Graf membantu kita menyatakan siapa berhubungan dengan siapa sebelum kita memilih model statistik yang sesuai.

Sebagai contoh, dalam studi penyebaran informasi, seseorang yang memiliki banyak koneksi mungkin lebih cepat menerima kabar. Di sini konsep derajat menjadi penting. Derajat sebuah simpul, ditulis $\deg(v)$, adalah banyaknya sisi yang bersisian dengan simpul v. Jika simpul adalah orang dan sisi adalah hubungan komunikasi, maka derajat seseorang adalah jumlah orang yang terhubung langsung dengannya. Konsep kecil ini nanti berkembang menjadi ukuran jaringan yang lebih kompleks, seperti centrality, clustering coefficient, dan community structure.

Ketiga, graf menghubungkan statistik dengan algoritma. Algoritma adalah prosedur langkah demi langkah yang dapat dijalankan untuk menyelesaikan masalah. Ketika data berbentuk graf, banyak pertanyaan statistik membutuhkan algoritma graf: mencari komponen terhubung, menghitung jarak terpendek, menemukan komunitas, membangun pohon merentang, melakukan traversal, atau menghitung ringkasan jaringan. Karena itu, buku ini tidak memisahkan definisi matematis dari implementasi komputasional. Sebuah definisi yang baik harus dapat dipakai untuk membuktikan teorema, tetapi juga harus dapat diterjemahkan menjadi struktur data dan algoritma.

Misalnya, graf dapat disimpan sebagai matriks ketetanggaan atau daftar ketetanggaan. Matriks ketetanggaan adalah matriks berukuran $n \times n$, dengan $n=|V|$, yang mencatat apakah dua simpul bertetangga. Daftar ketetanggaan menyimpan daftar tetangga untuk setiap simpul. Matriks memberi akses cepat untuk mengecek apakah dua simpul terhubung langsung, sedangkan daftar ketetanggaan sering lebih hemat memori untuk graf jarang, yaitu graf yang jumlah sisinya jauh lebih kecil daripada jumlah maksimum sisi yang mungkin. Perbandingan seperti ini adalah jembatan antara teori diskret dan praktik analisis data berskala besar.

Buku ini dimulai dari fondasi yang sangat sederhana karena di situlah ketelitian dibangun. Bab pertama memperkenalkan objek dasar graf: simpul, sisi, order, size, ketetanggaan, insidensi, subgraf, graf kosong, dan graf lengkap. Order sebuah graf adalah banyaknya simpul, yaitu $|V|$. Size sebuah graf adalah banyaknya sisi, yaitu $|E|$. Istilah-istilah ini tampak kecil, tetapi akan muncul berulang kali dalam pembuktian, algoritma, dan model.

Setelah itu, buku ini segera menekankan salah satu hukum paling dasar dalam teori graf, yaitu Handshaking Lemma:

$$\sum_{v\in V}\deg(v)=2|E|.$$

Maknanya sederhana: jika kita menjumlahkan semua derajat simpul dalam graf tak berarah, setiap sisi dihitung dua kali, sekali dari masing-masing ujungnya. Dari identitas ini diperoleh konsekuensi penting bahwa jumlah simpul berderajat ganjil selalu genap. Hasil ini merupakan contoh awal dari teknik penghitungan ganda, yaitu menghitung objek yang sama dengan dua cara berbeda. Teknik seperti ini sering muncul dalam teori graf dan kombinatorika, dan menjadi salah satu alat pembuktian utama dalam buku ini.

Untuk melihat intuisinya, ambil graf dengan sisi

$$E=\{\{A,B\},\{A,C\},\{B,D\},\{D,E\}\}.$$

Derajat tiap simpul adalah

$$\deg(A)=2,\quad \deg(B)=2,\quad \deg(C)=1,\quad \deg(D)=2,\quad \deg(E)=1.$$

Jumlah derajatnya adalah

$$2+2+1+2+1=8.$$

Karena jumlah sisi adalah |E|=4, maka

$$2|E|=8.$$

Kedua sisi persamaan cocok. Simpul berderajat ganjil adalah C dan E, jumlahnya dua, yaitu bilangan genap. Contoh kecil ini memperlihatkan pola umum yang akan terus kita gunakan: definisi yang tepat menghasilkan identitas yang kuat, dan identitas yang kuat menghasilkan konsekuensi struktural.

Setelah fondasi definisional, buku ini bergerak ke kelas-kelas graf. Kita akan membedakan graf sederhana, yaitu graf tanpa gelang dan tanpa sisi rangkap; graf berarah atau digraf, yaitu graf yang sisinya memiliki arah; dan graf berbobot, yaitu graf yang sisinya diberi nilai numerik. Nilai bobot dapat berarti jarak, biaya, risiko, waktu tempuh, kapasitas, frekuensi interaksi, atau kekuatan hubungan. Misalnya, dalam jaringan transportasi, bobot sisi dapat berupa waktu perjalanan antarstasiun. Dalam jaringan korelasi antarvariabel, bobot sisi dapat berupa ukuran asosiasi. Dalam jaringan risiko keuangan, bobot dapat merepresentasikan eksposur antarentitas.

Pada bagian tengah buku, kita akan mempelajari jalan, lintasan, siklus, keterhubungan, pohon, traversal, jarak terpendek, pohon merentang minimum, matching, pewarnaan, planaritas, dan aliran jaringan. Topik-topik ini membentuk inti teori graf klasik. Masing-masing akan diperlakukan bukan hanya sebagai daftar definisi, tetapi sebagai alat berpikir. Misalnya, konsep lintasan menjawab pertanyaan “apakah informasi dapat mengalir dari satu simpul ke simpul lain?” Konsep pohon menjawab pertanyaan “bagaimana menghubungkan semua simpul tanpa redundansi siklus?” Konsep aliran menjawab pertanyaan “berapa banyak kapasitas maksimum yang dapat dikirim dari sumber ke tujuan?”

Bagian berikutnya membawa graf ke wilayah aljabar linear dan spektral. Matriks ketetanggaan, matriks derajat, matriks insidensi, dan Laplacian graf akan menjadi cara untuk menghubungkan struktur diskret dengan ruang vektor, eigenvalue, dan eigenvector. Bagi pembaca statistik, bagian ini penting karena banyak metode modern—termasuk spectral clustering, reduksi dimensi berbasis graf, dan regularisasi pada jaringan—mengandalkan hubungan antara graf dan aljabar linear. Di sini graf tidak hanya dilihat sebagai kumpulan simpul dan sisi, tetapi juga sebagai operator linear yang menyimpan informasi struktural.

Setelah fondasi deterministik cukup kuat, buku ini masuk ke graf acak. Graf acak adalah graf yang dibangkitkan melalui mekanisme probabilistik. Dalam model Erdős–Rényi, misalnya, sisi-sisi dapat muncul secara acak dengan probabilitas tertentu. Model seperti ini memberi cara untuk bertanya: apakah pola jaringan yang kita amati tampak luar biasa, atau dapat muncul secara wajar karena peluang acak? Teori graf acak menjadi salah satu dasar penting bagi analisis jaringan modern; perlakuan klasik dan mendalam dapat ditemukan dalam Bollobás (2001).

Dari graf acak, kita bergerak ke model statistik untuk jaringan. Di sini graf bukan hanya struktur pembantu, tetapi objek data itu sendiri. Kita akan memperkenalkan model seperti stochastic block model, degree-corrected block model, latent space model, dan exponential random graph model. Tujuan statistiknya meliputi estimasi parameter, deteksi komunitas, pengujian hipotesis, validasi model, dan interpretasi ketidakpastian. Pembaca akan diajak membedakan ringkasan deskriptif, seperti distribusi derajat, dari inferensi statistik, seperti klaim bahwa suatu jaringan memiliki struktur komunitas yang lebih kuat daripada yang diharapkan di bawah model acak tertentu.

Bagian akhir buku menghubungkan teori graf dengan model grafis probabilistik dan kausalitas. Dalam model grafis probabilistik, graf menyatakan hubungan dependensi antarvariabel acak. Dalam kausalitas, graf berarah asiklik membantu merumuskan asumsi tentang sebab dan akibat. Di sini kehati-hatian sangat penting. Sisi dalam graf tidak selalu berarti sebab-akibat. Sebuah sisi dapat berarti kedekatan, asosiasi, interaksi, ketergantungan bersyarat, atau relasi struktural lain, tergantung definisi model. Salah satu tujuan buku ini adalah melatih pembaca untuk selalu bertanya: “Apa arti simpul? Apa arti sisi? Apa asumsi yang membuat interpretasi ini sah?”

Karena buku ini ditujukan untuk tingkat pascasarjana, pembaca diharapkan bersedia bekerja dengan definisi formal, bukti, dan algoritma. Namun formal tidak berarti kering. Definisi formal justru membantu kita menghindari ambiguitas. Bukti membantu kita mengetahui mengapa suatu klaim benar, bukan sekadar percaya bahwa klaim itu benar. Algoritma membantu kita mengubah gagasan matematis menjadi prosedur yang dapat dijalankan pada data nyata.

Sikap belajar yang disarankan adalah sebagai berikut. Ketika bertemu definisi, buatlah contoh dan bukan-contoh. Ketika bertemu teorema, tanyakan asumsi apa yang diperlukan. Ketika bertemu algoritma, cari invariant-nya, yaitu sifat yang tetap benar selama algoritma berjalan. Ketika bertemu model statistik berbasis graf, tanyakan sumber ketidakpastiannya: apakah ketidakpastian berasal dari sampling, dari pembentukan sisi, dari parameter model, atau dari proses generatif yang diasumsikan?

Sebagai contoh, jika kita mendefinisikan graf sederhana sebagai graf tanpa gelang dan tanpa sisi rangkap, maka graf dengan sisi A,A\ bukan graf sederhana karena A,A\ adalah gelang, yaitu sisi yang menghubungkan simpul ke dirinya sendiri. Graf dengan dua sisi berbeda yang sama-sama menghubungkan A dan B juga bukan graf sederhana karena memiliki sisi rangkap. Contoh dan bukan-contoh seperti ini membuat definisi menjadi tajam.

Jika kita mempelajari algoritma BFS, jangan hanya menghafal bahwa BFS mencari simpul “melebar dulu”. Tanyakan: struktur data apa yang dipakai? Simpul mana yang sudah dikunjungi? Mengapa jarak yang ditemukan BFS pada graf tak berbobot adalah jarak terpendek dalam jumlah sisi? Pertanyaan seperti ini mengubah algoritma dari resep menjadi pengetahuan.

Jika kita mempelajari Bayesian network, jangan hanya menggambar panah antarvariabel. Tanyakan: apakah grafnya asiklik? Apa arti panah dalam model? Independensi bersyarat apa yang dinyatakan? Apakah klaim kausal memerlukan asumsi tambahan? Dengan cara ini, teori graf menjadi alat statistik yang bertanggung jawab, bukan sekadar visualisasi yang meyakinkan.

Buku ini memiliki satu garis besar: dari objek diskret yang paling sederhana menuju model statistik yang menangani ketidakpastian dan dependensi. Kita mulai dari G=(V,E), lalu membangun derajat, pembuktian, representasi, algoritma, struktur, optimisasi, matriks, graf acak, jaringan, model grafis probabilistik, dan kausalitas. Setiap bagian bergantung pada bagian sebelumnya. Karena itu, fondasi awal tidak boleh dilewati terlalu cepat.

Pada akhirnya, tujuan buku ini bukan hanya agar pembaca dapat menyebutkan definisi graf, tetapi agar pembaca dapat memakai graf secara benar dalam penelitian statistik. Pembaca diharapkan mampu membangun graf dari data, memilih representasi yang sesuai, membuktikan klaim dasar, menjalankan algoritma, memahami keterbatasan komputasi, memodelkan jaringan secara probabilistik, dan membedakan interpretasi deskriptif, prediktif, serta kausal.

Teori graf tampak sederhana karena dimulai dari titik dan garis. Tetapi dari titik dan garis itu muncul bahasa yang sangat luas untuk memahami hubungan. Dalam statistik, hubungan adalah pusat dari banyak pertanyaan ilmiah: siapa memengaruhi siapa, variabel mana terkait dengan variabel mana, struktur apa yang tersembunyi dalam data, dan seberapa yakin kita terhadap kesimpulan itu. Buku ini adalah perjalanan untuk menjawab pertanyaan-pertanyaan tersebut dengan definisi yang jelas, bukti yang hati-hati, algoritma yang dapat dijalankan, dan pemodelan statistik yang sadar akan ketidakpastian.

References

Bollobás, B. (2001). Random Graphs (2nd ed.). Cambridge University Press.

Diestel, R. (2017). Graph Theory (5th ed.). Springer.

Kolaczyk, E. D. (2009). Statistical Analysis of Network Data: Methods and Models. Springer.

Lauritzen, S. L. (1996). Graphical Models. Oxford University Press.

Newman, M. E. J. (2018). Networks (2nd ed.). Oxford University Press.

Pearl, J. (2009). Causality: Models, Reasoning, and Inference (2nd ed.). Cambridge University Press.

West, D. B. (2001). Introduction to Graph Theory (2nd ed.). Prentice Hall.

Document information

Pendahuluan

Project	Teori Graf untuk Statistik
Document	Document 1.4
Author	Harizahayu
Verifier	Not verified
Downloaded	July 08, 2026 10:33 KST
Status	Working
Document link	https://theorytrace.com/projects/teori-graf-untuk-statistik/documents/pendahuluan/