Bab 1: Objek Dasar Graf

Statistik sering mengajarkan kita untuk bertanya: apa unit pengamatan kita, apa variabelnya, dan bagaimana nilai-nilainya tersebar? Teori graf menambahkan satu pertanyaan yang sama pentingnya: objek mana berhubungan dengan objek mana?

Jika data biasa berbentuk tabel, maka graf adalah bentuk data untuk relasi. Dalam tabel survei, satu baris dapat mewakili satu responden. Dalam graf, satu simpul dapat mewakili satu responden, dan satu sisi dapat mewakili hubungan antara dua responden. Hubungan itu dapat berupa pertemanan, komunikasi, transaksi, konflik jadwal, kedekatan geografis, korelasi tinggi, atau ketergantungan antarvariabel.

Bab ini membangun bahasa dasar tersebut secara formal. Kita akan mulai dari definisi graf sebagai

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

lalu memperjelas istilah-istilah pertama: simpul, sisi, order, size, ketetanggaan, insidensi, derajat, subgraf, graf kosong, dan graf lengkap. Definisi-definisi ini tampak sederhana, tetapi hampir seluruh teori graf dan algoritma graf berdiri di atasnya. Notasi dan terminologi yang digunakan di sini mengikuti konvensi umum dalam buku-buku teori graf standar seperti Diestel dan West, dengan penyesuaian bahasa untuk konteks statistik (Diestel, 2017; West, 2001).

1.1 Graf sebagai pasangan G=(V,E)

Secara formal, sebuah graf adalah pasangan

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

dengan:

- V adalah himpunan simpul, disebut juga vertices atau nodes;
- E adalah himpunan sisi, disebut juga edges;
- dalam buku ini, kecuali dinyatakan lain, V dianggap tak kosong.

Simpul adalah objek dasar yang sedang kita pelajari. Sisi adalah relasi antara objek-objek tersebut.

Untuk graf tak berarah sederhana, sebuah sisi biasanya ditulis sebagai pasangan tak berurutan:

$$\{u, v\},$$

dengan u,v ∈ V dan u ≠ v. Kata “tak berurutan” berarti

$$\{u, v\} = \{v, u\}.$$

Jadi, sisi A,B\ sama dengan sisi B,A\. Ini cocok untuk relasi yang bersifat simetris, misalnya “berteman dengan”, “berbatasan dengan”, atau “pernah berada dalam satu kelompok eksperimen”.

Contoh 1.1: Graf pertemanan sederhana

Misalkan terdapat empat responden survei:

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

Kita mencatat relasi “saling mengenal” sebagai berikut:

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

Maka grafnya adalah

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

Interpretasinya:

- A mengenal B;

- A mengenal C;

- B mengenal D;

- tidak ada sisi antara C dan D, sehingga data ini tidak mencatat bahwa C mengenal D.

Perhatikan bahwa tidak adanya sisi bukan selalu berarti relasi itu “pasti tidak ada” di dunia nyata. Dalam statistik, data dapat tidak lengkap. Jadi, secara matematis kita mengatakan “tidak ada sisi dalam graf yang diberikan”, sedangkan secara substantif kita harus bertanya apakah ketiadaan sisi berarti ketiadaan relasi, ketiadaan observasi, atau hasil penyaringan tertentu. Isu seperti ini penting dalam analisis jaringan statistik, karena data jaringan sering dipengaruhi oleh desain pengambilan sampel, definisi relasi, dan batas populasi yang diamati (Wasserman & Faust, 1994; Kolaczyk & Csárdi, 2014).

1.2 Simpul dan sisi

Istilah simpul dan sisi perlu dipahami dengan sangat konkret.

Sebuah simpul adalah satu entitas yang dianggap sebagai unit dalam graf. Dalam aplikasi statistik, simpul dapat berupa:

- responden survei;
- rumah tangga;
- sekolah;
- wilayah geografis;
- variabel acak;
- gen;
- dokumen;
- akun pengguna;
- titik waktu;
- kelompok perlakuan dalam eksperimen.

Sebuah sisi adalah relasi antara dua simpul. Sisi dapat menyatakan:

- dua orang saling mengenal;
- dua wilayah berbatasan;
- dua variabel memiliki hubungan dependensi;
- dua gen berinteraksi;
- dua dokumen memiliki kemiripan tinggi;
- dua unit tidak boleh ditempatkan pada kelompok yang sama;
- satu simpul dapat dicapai dari simpul lain dalam suatu jaringan.

Pada bab ini kita fokus terlebih dahulu pada graf tak berarah sederhana. Artinya, sisi tidak mempunyai arah dan tidak ada dua sisi rangkap antara pasangan simpul yang sama. Kelas-kelas graf yang lebih umum—seperti multigraf, graf berarah, dan graf berbobot—akan dibahas lebih sistematis pada Bab 3.

Contoh 1.2: Graf wilayah bertetangga

Misalkan kita memiliki empat wilayah:

$$V = \{\text{Utara}, \text{Timur}, \text{Selatan}, \text{Barat}\}.$$

Jika dua wilayah berbagi perbatasan, kita hubungkan dengan sisi. Misalnya,

$$E = \{\{\text{Utara}, \text{Timur}\}, \{\text{Timur}, \text{Selatan}\}, \{\text{Selatan}, \text{Barat}\}, \{\text{Barat}, \text{Utara}\}\}.$$

Graf ini dapat digunakan sebagai struktur awal dalam analisis spasial. Misalnya, ketika kita memodelkan angka penyakit per wilayah, kita mungkin ingin mencatat wilayah mana yang bertetangga karena nilai di wilayah berdekatan sering dipelajari bersama dalam model spasial. Bab ini belum membahas model statistiknya, tetapi graf sudah menyediakan bahasa untuk menyatakan struktur kedekatan.

1.3 Order dan size

Setelah graf didefinisikan, dua ukuran paling awal adalah order dan size.

Order dari graf G=(V,E) adalah banyaknya simpul dalam graf tersebut:

$$|V|.$$

Size dari graf G=(V,E) adalah banyaknya sisi dalam graf tersebut:

$$|E|.$$

Notasi |S| berarti banyaknya elemen dalam himpunan S. Jadi, jika V berisi 5 simpul, maka |V|=5. Jika E berisi 7 sisi, maka |E|=7.

Dalam literatur teori graf, order biasanya merujuk pada jumlah simpul, sedangkan size merujuk pada jumlah sisi (West, 2001; Diestel, 2017).

Contoh 1.3: Menghitung order dan size

Ambil graf

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

dan

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

Maka:

$$|V| = 5,$$

sehingga order graf adalah 5.

Sementara itu,

$$|E| = 4,$$

sehingga size graf adalah 4.

Secara substantif, jika graf ini adalah graf pertemanan, maka terdapat 5 orang dan 4 hubungan pertemanan yang tercatat.

Order dan size dalam data statistik

Dalam analisis jaringan, order sering berkaitan dengan ukuran populasi atau jumlah unit yang dianalisis. Size berkaitan dengan jumlah relasi yang teramati. Namun, dua graf dengan order dan size yang sama belum tentu memiliki struktur yang sama.

Sebagai contoh, dua graf berikut sama-sama memiliki 4 simpul dan 3 sisi:

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

Graf pertama:

$$E_1 = \{\{A,B\},\{B,C\},\{C,D\}\}.$$

Graf kedua:

$$E_2 = \{\{A,B\},\{A,C\},\{A,D\}\}.$$

Keduanya memiliki order 4 dan size 3. Tetapi strukturnya berbeda. Graf pertama membentuk rantai:

$$A - B - C - D,$$

sedangkan graf kedua memiliki satu simpul pusat A yang terhubung ke semua simpul lain. Dalam analisis sosial, graf kedua dapat menunjukkan aktor yang lebih sentral. Dalam analisis komputasi, struktur seperti ini dapat memengaruhi jarak, traversal, dan penyebaran informasi.

Jadi, order dan size adalah ringkasan awal, bukan deskripsi lengkap.

1.4 Ketetanggaan

Dua simpul u dan v disebut bertetangga atau adjacent jika ada sisi yang menghubungkan keduanya.

Dalam graf tak berarah sederhana,

$$u \text{ bertetangga dengan } v$$

jika dan hanya jika

$$\{u,v\} \in E.$$

Kata “jika dan hanya jika” berarti dua arah sekaligus:

1. jika u,v\∈ E, maka u dan v bertetangga;
2. jika u dan v bertetangga, maka u,v\∈ E.

Notasi “jika dan hanya jika” sering disingkat sebagai “iff” dalam literatur berbahasa Inggris, atau ditulis dengan simbol Longleftrightarrow.

Maka definisi ketetanggaan dapat ditulis:

$$u \sim v \iff \{u, v\} \in E.$$

Simbol u sim v dibaca “u bertetangga dengan v”.

Contoh 1.4: Membaca ketetanggaan

Misalkan

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

dan

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

Maka:

- A bertetangga dengan B, karena A,B\∈ E;
- A bertetangga dengan D, karena A,D\∈ E;
- B bertetangga dengan C, karena B,C\∈ E;
- A tidak bertetangga dengan C, karena A,C\∉ E;
- C tidak bertetangga dengan D, karena C,D\∉ E.

Dalam graf tak berarah, ketetanggaan bersifat simetris:

$$u \sim v \iff v \sim u.$$

Jika A bertetangga dengan B, maka B bertetangga dengan A. Ini berbeda dari graf berarah, yang akan kita bahas kemudian. Dalam graf berarah, relasi “mengikuti akun” atau “mengirim pesan kepada” tidak harus simetris.

1.5 Insidensi

Istilah berikutnya adalah insidensi.

Sebuah sisi e disebut bersisian atau incident dengan simpul v jika simpul v merupakan salah satu ujung dari sisi e.

Dalam graf tak berarah sederhana, jika

$$e = \{u, v\},$$

maka sisi e bersisian dengan u dan juga bersisian dengan v.

Ketetanggaan adalah relasi antara simpul dan simpul. Insidensi adalah relasi antara simpul dan sisi.

Contoh 1.5: Ketetanggaan versus insidensi

Misalkan

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

Sisi

$$e_1 = \{A, B\}$$

bersisian dengan A dan B, tetapi tidak bersisian dengan C atau D.

Simpul A bertetangga dengan B, karena ada sisi A,B\. Simpul A juga bertetangga dengan C, karena ada sisi A,C\.

Namun, sisi A,B\ tidak “bertetangga” dengan A. Istilah yang tepat adalah sisi A,B\ bersisian dengan A.

Perbedaan kecil ini penting karena banyak definisi graf memakai bahasa yang presisi. Misalnya, derajat simpul didefinisikan melalui banyaknya sisi yang bersisian dengan simpul tersebut.

1.6 Derajat simpul

Derajat sebuah simpul v, ditulis

$$\deg(v),$$

adalah banyaknya sisi yang bersisian dengan v.

Untuk graf tak berarah sederhana, derajat v juga sama dengan banyaknya tetangga dari v. Kesetaraan ini benar karena dalam graf sederhana tidak ada sisi rangkap dan tidak ada gelang. Pada kelas graf yang lebih umum, terutama multigraf dan graf dengan gelang, definisi derajat perlu diperlakukan lebih hati-hati. Bab 2 akan membahas derajat secara lebih dalam dan membuktikan Handshaking Lemma.

Contoh 1.6: Menghitung derajat

Misalkan

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

dan

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

Mari hitung derajat setiap simpul.

Simpul A bersisian dengan tiga sisi:

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

Jadi,

$$\deg(A) = 3.$$

Simpul B hanya bersisian dengan A,B\, sehingga

$$\deg(B) = 1.$$

Simpul C hanya bersisian dengan A,C\, sehingga

$$\deg(C) = 1.$$

Simpul D bersisian dengan A,D\ dan D,E\, sehingga

$$\deg(D) = 2.$$

Simpul E hanya bersisian dengan D,E\, sehingga

$$\deg(E) = 1.$$

Ringkasnya:

$$\deg(A) = 3, \quad \deg(B) = 1, \quad \deg(C) = 1, \quad \deg(D) = 2, \quad \deg(E) = 1.$$

Dalam konteks jaringan sosial, derajat dapat dibaca sebagai jumlah relasi langsung yang dimiliki aktor. Dalam jaringan komunikasi, derajat dapat menjadi jumlah kontak langsung. Namun, interpretasi statistik derajat harus bergantung pada bagaimana sisi didefinisikan dan bagaimana data dikumpulkan. Buku-buku analisis jaringan menekankan bahwa ukuran seperti derajat adalah statistik deskriptif yang berguna, tetapi tidak otomatis menjadi bukti kausal atau bukti kepentingan substantif tanpa konteks pengukuran yang tepat (Wasserman & Faust, 1994; Kolaczyk & Csárdi, 2014).

1.7 Lingkungan sebuah simpul

Setelah memahami ketetanggaan dan derajat, kita dapat mendefinisikan lingkungan atau neighborhood dari sebuah simpul.

Untuk graf tak berarah sederhana G=(V,E), lingkungan terbuka dari simpul v, ditulis

$$N(v),$$

adalah himpunan semua simpul yang bertetangga dengan v:

$$N(v) = \{u \in V : \{u, v\} \in E\}.$$

Karena N(v) adalah himpunan tetangga v, maka pada graf sederhana berlaku

$$\deg(v) = |N(v)|.$$

Contoh 1.7: Lingkungan simpul

Misalkan

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

dan

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

Tetangga A adalah B,C,D. Jadi,

$$N(A) = \{B, C, D\}.$$

Tetangga D adalah A dan E. Jadi,

$$N(D) = \{A, E\}.$$

Tetangga B hanya A. Jadi,

$$N(B) = \{A\}.$$

Karena |N(A)|=3, maka deg(A)=3. Karena |N(D)|=2, maka deg(D)=2.

Lingkungan simpul sering menjadi objek penting dalam statistik jaringan. Misalnya, jika simpul adalah individu, maka N(v) dapat mewakili kontak langsung individu v. Jika simpul adalah variabel, maka N(v) dapat mewakili variabel-variabel yang dihubungkan langsung dengan v menurut kriteria tertentu.

1.8 Graf sebagai model: apa yang dipertahankan dan apa yang diabaikan?

Graf adalah model abstrak. Artinya, graf mempertahankan sebagian informasi dan mengabaikan sebagian informasi lain.

Misalnya, dalam graf pertemanan sederhana, kita hanya mencatat apakah dua orang berteman atau tidak. Kita mengabaikan:

- sejak kapan mereka berteman;
- seberapa sering mereka berkomunikasi;
- seberapa kuat hubungan mereka;
- apakah hubungan itu simetris;
- apakah data pertemanan diperoleh dari satu pihak atau dua pihak.

Jika informasi kekuatan relasi penting, kita mungkin memakai graf berbobot. Jika arah relasi penting, kita mungkin memakai graf berarah. Jika dua orang dapat memiliki lebih dari satu jenis hubungan, kita mungkin memakai multigraf atau graf berlapis. Untuk saat ini, kita mulai dari bentuk paling dasar agar struktur matematisnya jelas.

Contoh 1.8: Korelasi sebagai graf

Misalkan kita mengukur empat variabel:

$$V = \{X_1, X_2, X_3, X_4\}.$$

Kita dapat membuat sisi antara dua variabel jika korelasi sampelnya melebihi ambang tertentu, misalnya $|r|>0.7$. Jika hasilnya adalah

$$E = \{\{X_1, X_2\}, \{X_2, X_3\}\},$$

maka graf menyatakan bahwa X_1 sangat berkorelasi dengan X_2, dan X_2 sangat berkorelasi dengan X_3, menurut aturan ambang yang dipilih.

Namun, graf ini tidak menyimpan nilai korelasi persisnya. Relasi dengan $r=0.71$ dan $r=0.95$ sama-sama menjadi sisi. Jadi, graf sederhana mengubah informasi numerik menjadi informasi diskret: ada sisi atau tidak ada sisi.

Ini berguna ketika kita ingin mempelajari pola relasi secara struktural, tetapi harus dilakukan dengan hati-hati. Pemilihan ambang dapat mengubah graf secara besar. Dalam analisis statistik jaringan, cara membangun graf dari data adalah bagian penting dari pemodelan, bukan langkah teknis yang netral (Kolaczyk & Csárdi, 2014).

1.9 Subgraf

Sering kali kita tidak ingin melihat seluruh graf sekaligus. Kita mungkin ingin mempelajari sebagian simpul, sebagian sisi, atau struktur lokal di sekitar kelompok tertentu. Untuk itu kita memakai konsep subgraf.

Misalkan

$$G = (V, E)$$

adalah graf. Graf

$$H = (W, F)$$

disebut subgraf dari G jika:

$$W \subseteq V$$

dan

$$F \subseteq E,$$

dengan setiap sisi dalam F hanya menghubungkan simpul-simpul yang ada di W.

Dengan kata lain, subgraf diperoleh dengan mengambil sebagian simpul dan sebagian sisi dari graf asal, tanpa menciptakan sisi baru yang sebelumnya tidak ada.

Notasi yang sering dipakai adalah

$$H \subseteq G.$$

Definisi subgraf adalah salah satu definisi dasar dalam teori graf dan muncul dalam berbagai pembahasan seperti keterhubungan, pohon merentang, clique, matching, dan minor graf (Diestel, 2017; West, 2001).

Contoh 1.9: Subgraf dari jaringan responden

Misalkan graf besar

$$G = (V, E)$$

memiliki

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

dan

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

Ambil

$$W = \{A, B, D\}.$$

Kita dapat membentuk subgraf

$$H = (W, F)$$

dengan

$$F = \{\{A, B\}, \{B, D\}\}.$$

Ini valid karena:

- W⊆ V;

- F⊆ E;

- setiap sisi dalam F menghubungkan simpul-simpul di W.

Tetapi kita tidak boleh memasukkan sisi A,D\ ke dalam F, karena A,D\∉ E. Subgraf tidak boleh menciptakan relasi baru.

1.10 Subgraf terinduksi

Ada jenis subgraf yang sangat penting: subgraf terinduksi.

Misalkan G=(V,E) adalah graf dan W⊆ V. Subgraf yang diinduksi oleh W, ditulis

$$G[W],$$

adalah graf dengan himpunan simpul W dan semua sisi dari G yang kedua ujungnya berada di W.

Secara formal,

$$G[W] = (W, F),$$

dengan

$$F = \{\{u, v\} \in E : u \in W,\ v \in W\}.$$

Perbedaan pentingnya adalah:

- subgraf biasa boleh memilih sebagian sisi;

- subgraf terinduksi harus mengambil semua sisi yang tersedia di antara simpul-simpul yang dipilih.

Contoh 1.10: Subgraf terinduksi

Gunakan graf sebelumnya:

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

dan

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

Ambil

$$W = \{A, B, D\}.$$

Sisi-sisi dalam E yang kedua ujungnya berada di W adalah:

$$\{A,B\},\{B,D\}.$$

Maka

$$G[W]=(\{A,B,D\},\{\{A,B\},\{B,D\}\}).$$

Sekarang ambil

$$W'=\{A,B,C,D\}.$$

Sisi-sisi yang kedua ujungnya berada di W' adalah:

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

Jadi,

$$G[W']=(\{A,B,C,D\},\{\{A,B\},\{A,C\},\{B,D\},\{C,D\}\}).$$

Subgraf terinduksi penting dalam statistik karena kita sering mengambil subset unit pengamatan. Misalnya, kita mungkin ingin melihat jaringan hanya untuk responden dari satu wilayah, satu kelompok umur, atau satu kelompok perlakuan. Jika kita mengambil semua relasi yang ada di antara anggota subset itu, kita sedang membentuk subgraf terinduksi.

1.11 Graf kosong dan graf trivial

Istilah “graf kosong” dapat dipakai berbeda-beda di beberapa teks. Karena itu, kita perlu menetapkan konvensi.

Dalam buku ini, karena graf didefinisikan dengan V tak kosong, graf kosong pada himpunan simpul V berarti graf yang tidak memiliki sisi:

$$G=(V,\varnothing).$$

Di sini varnothing adalah himpunan kosong, yaitu himpunan yang tidak memiliki elemen.

Jadi graf kosong dalam konvensi ini memiliki simpul, tetapi tidak memiliki sisi. Kadang-kadang graf seperti ini juga disebut graf tanpa sisi atau edgeless graph. Beberapa literatur menggunakan istilah null graph untuk graf dengan himpunan simpul kosong, tetapi konvensi itu tidak kita pakai sebagai objek utama karena sejak awal V kita anggap tak kosong. Perbedaan terminologi seperti ini memang ada dalam literatur, sehingga selalu baik memeriksa definisi yang dipakai oleh suatu buku atau artikel (West, 2001).

Contoh 1.11: Graf kosong pada empat simpul

Misalkan

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

Graf kosong pada V adalah

$$G = (V, \varnothing).$$

Graf ini memiliki:

$$|V| = 4$$

dan

$$|E| = 0.$$

Setiap simpul memiliki derajat 0:

$$\deg(A) = \deg(B) = \deg(C) = \deg(D) = 0.$$

Dalam konteks data, ini dapat berarti tidak ada relasi yang tercatat di antara empat unit. Tetapi lagi-lagi interpretasi substantifnya perlu hati-hati. Mungkin memang tidak ada relasi, atau mungkin relasi belum diukur.

Graf trivial

Graf dengan tepat satu simpul dan tidak ada sisi disebut graf trivial.

Jika

$$V = \{A\}$$

dan

$$E = \varnothing,$$

maka

$$G = (\{A\}, \varnothing)$$

adalah graf trivial.

Graf trivial memiliki order 1 dan size 0. Walaupun terlihat terlalu sederhana, graf trivial sering muncul sebagai kasus dasar dalam pembuktian induksi. Misalnya, ketika kelak kita membuktikan sifat pohon, kasus satu simpul sering menjadi titik awal yang penting.

1.12 Graf lengkap

Di ujung lain dari graf kosong terdapat graf lengkap.

Graf lengkap pada n simpul, ditulis

$$K_n,$$

adalah graf sederhana dengan n simpul di mana setiap pasangan simpul yang berbeda bertetangga.

Artinya, jika

$$V = \{v_1, v_2, \ldots, v_n\},$$

maka untuk setiap i≠ j, sisi

$$\{v_i, v_j\}$$

ada di E.

Notasi K_n adalah notasi standar untuk graf lengkap pada n simpul (Diestel, 2017; West, 2001).

Contoh 1.12: Graf lengkap K_4

Misalkan

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

Graf lengkap pada empat simpul, K_4, memiliki semua sisi antara pasangan simpul berbeda:

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

Jadi K_4 memiliki order 4 dan size 6.

Setiap simpul bertetangga dengan tiga simpul lain. Maka:

$$\deg(A) = \deg(B) = \deg(C) = \deg(D) = 3.$$

Secara statistik, K_4 dapat muncul sebagai graf konflik jika setiap pasangan unit tidak boleh ditempatkan bersama, atau sebagai graf pertemanan jika setiap orang dalam kelompok saling mengenal. Namun, graf lengkap sering lebih berguna sebagai objek pembanding teoretis: ia merepresentasikan relasi maksimum yang mungkin pada n simpul dalam graf sederhana.

1.13 Banyak sisi pada graf lengkap

Untuk graf lengkap K_n, berapa banyak sisi yang dimiliki?

Karena setiap sisi menghubungkan sepasang simpul berbeda, maka jumlah sisi sama dengan banyaknya cara memilih 2 simpul dari n simpul.

Jumlah cara memilih 2 objek dari n objek adalah

$$\binom{n}{2}.$$

Rumusnya:

$$\binom{n}{2} = \frac{n(n-1)}{2}.$$

Jadi,

$$|E(K_n)| = \binom{n}{2} = \frac{n(n-1)}{2}.$$

Mengapa dibagi 2?

Jika kita menghitung secara berurutan, setiap simpul dapat dipasangkan dengan n-1 simpul lain. Maka tampaknya ada

$$n(n-1)$$

pasangan.

Tetapi pasangan (u,v) dan (v,u) menyatakan sisi tak berarah yang sama. Jadi setiap sisi terhitung dua kali. Karena itu kita bagi 2:

$$|E(K_n)| = \frac{n(n-1)}{2}.$$

Contoh 1.13: Size dari K_5

Untuk K_5,

$$|E(K_5)| = \frac{5(5-1)}{2} = \frac{5 \cdot 4}{2} = 10.$$

Jadi graf lengkap pada 5 simpul memiliki 10 sisi.

Jika lima variabel semuanya dihubungkan satu sama lain, maka graf hubungan penuh antarvariabel memiliki 10 pasangan variabel. Dalam statistik multivariat, jumlah pasangan variabel tumbuh kuadratik terhadap jumlah variabel. Ini menjadi salah satu alasan mengapa struktur graf dapat membantu meringkas atau membatasi relasi yang perlu dipertimbangkan.

1.14 Graf sederhana sebagai ruang kerja awal

Sampai titik ini, graf yang kita gunakan adalah graf tak berarah sederhana. Secara formal, graf sederhana dapat dipahami sebagai

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

dengan

$$E \subseteq \{\{u, v\} : u, v \in V,\ u \neq v\}.$$

Artinya, setiap sisi adalah himpunan berisi dua simpul berbeda. Definisi ini otomatis mengecualikan dua hal:

1. gelang atau loop, yaitu sisi dari simpul ke dirinya sendiri, misalnya A,A\;
2. sisi rangkap, yaitu lebih dari satu sisi antara pasangan simpul yang sama.

Karena E adalah himpunan, elemen yang sama tidak dapat muncul dua kali. Maka dalam graf sederhana tidak mungkin ada dua salinan berbeda dari sisi A,B\.

Contoh 1.14: Apa yang bukan graf sederhana?

Misalkan seseorang menulis:

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

Sebagai himpunan biasa, penulisan ini sebenarnya sama dengan

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

karena elemen duplikat tidak dihitung dua kali. Jika kita ingin memperbolehkan dua sisi berbeda antara A dan B, kita perlu memakai struktur yang lebih umum, yaitu multigraf.

Sekarang lihat sisi

$$\{A, A\}.$$

Sisi ini menghubungkan A dengan dirinya sendiri. Dalam graf sederhana, sisi seperti ini tidak diperbolehkan. Sisi semacam itu disebut gelang atau loop.

Mengapa kita mulai dari graf sederhana? Karena sebagian besar konsep dasar—ketetanggaan, derajat, subgraf, graf lengkap, jalan, lintasan, keterhubungan—lebih mudah dipahami tanpa komplikasi gelang dan sisi rangkap. Setelah fondasinya kuat, generalisasi ke graf lain menjadi lebih masuk akal.

1.15 Contoh statistik pertama: graf konflik eksperimen

Graf dapat membantu merancang atau memeriksa struktur eksperimen. Misalkan kita memiliki lima unit eksperimen:

$$V = \{U_1, U_2, U_3, U_4, U_5\}.$$

Dua unit dihubungkan oleh sisi jika keduanya tidak boleh menerima perlakuan yang sama. Misalnya, mungkin karena dua unit berada terlalu dekat secara spasial, berasal dari rumah tangga yang sama, atau saling memengaruhi.

Misalkan

$$E = \{\{U_1, U_2\}, \{U_1, U_3\}, \{U_2, U_4\}, \{U_3, U_5\}\}.$$

Graf ini disebut graf konflik dalam konteks pemodelan konflik diskret. Ia belum menyelesaikan masalah desain eksperimen, tetapi ia mengubah kendala substantif menjadi objek matematis. Kelak, pada bab pewarnaan graf, kita akan melihat bahwa memberi warna pada simpul dapat ditafsirkan sebagai menugaskan kelompok atau jadwal sedemikian rupa sehingga simpul bertetangga tidak mendapat warna yang sama.

Untuk saat ini, yang perlu diperhatikan adalah:

- simpul = unit eksperimen;
- sisi = konflik atau larangan ditempatkan bersama;
- derajat U_i = banyaknya unit lain yang berkonflik langsung dengan U_i.

Hitung derajatnya:

$$\deg(U_1) = 2,$$

karena U_1 berkonflik dengan U_2 dan U_3.

$$\deg(U_2) = 2,$$

karena U_2 berkonflik dengan U_1 dan U_4.

$$\deg(U_3) = 2,$$

karena U_3 berkonflik dengan U_1 dan U_5.

$$\deg(U_4) = 1, \quad \deg(U_5) = 1.$$

Graf ini memiliki order 5 dan size 4.

1.16 Contoh statistik kedua: graf kemiripan observasi

Dalam analisis data, kita kadang membangun graf dari ukuran kemiripan. Misalkan ada enam observasi:

$$V = \{1, 2, 3, 4, 5, 6\}.$$

Setiap observasi memiliki vektor fitur, misalnya hasil pengukuran laboratorium, skor survei, atau representasi numerik dokumen. Kita hubungkan dua observasi dengan sisi jika jaraknya lebih kecil dari ambang tertentu.

Misalkan setelah menghitung jarak, kita memperoleh

$$E = \{\{1, 2\}, \{1, 3\}, \{2, 3\}, \{4, 5\}, \{5, 6\}\}.$$

Graf ini mengatakan bahwa observasi 1,2,3 saling dekat, sedangkan 4,5,6 membentuk struktur lain dengan sisi \4,5\ dan \5,6\.

Derajatnya:

$$\deg(1) = 2, \quad \deg(2) = 2, \quad \deg(3) = 2,$$

$$\deg(4) = 1, \quad \deg(5) = 2, \quad \deg(6) = 1.$$

Subgraf terinduksi pada

$$W = \{1, 2, 3\}$$

adalah graf lengkap K_3, karena semua pasangan di antara 1,2,3 terhubung.

Subgraf terinduksi pada

$$W' = \{4, 5, 6\}$$

bukan graf lengkap, karena sisi \4,6\ tidak ada.

Contoh ini memperlihatkan bagaimana struktur graf dapat memunculkan pola kelompok. Namun, pada tahap ini kita belum melakukan inferensi. Kita baru membangun representasi diskret dari relasi kemiripan.

1.17 Contoh statistik ketiga: graf variabel

Sekarang misalkan simpul bukan individu, melainkan variabel.

Ambil lima variabel:

$$V = \{X_1, X_2, X_3, X_4, X_5\}.$$

Kita membuat sisi jika dua variabel memiliki hubungan yang dianggap cukup kuat menurut kriteria tertentu. Misalnya, sisi dapat dibuat jika korelasi absolut lebih dari 0.6, atau jika ada alasan teoritis bahwa dua variabel berhubungan langsung.

Misalkan

$$E = \{\{X_1, X_2\}, \{X_2, X_3\}, \{X_2, X_4\}, \{X_4, X_5\}\}.$$

Maka:

$$N(X_2) = \{X_1, X_3, X_4\},$$

sehingga

$$\deg(X_2) = 3.$$

Sementara itu,

$$N(X_5) = \{X_4\},$$

sehingga

$$\deg(X_5) = 1.$$

Graf seperti ini dapat menjadi langkah awal untuk mempelajari struktur dependensi antarvariabel. Namun, penting untuk tidak menyamakan sisi dengan kausalitas. Sisi yang dibuat dari korelasi hanya menyatakan hubungan berdasarkan kriteria korelasi yang dipilih. Klaim kausal memerlukan asumsi dan desain analisis yang jauh lebih kuat. Buku ini akan kembali ke isu tersebut ketika membahas DAG dan kausalitas pada bab-bab akhir.

1.18 Kesalahan umum pada tahap definisi

Sebelum menutup bab, mari kita periksa beberapa kesalahan umum.

Kesalahan 1: Menganggap sisi selalu berarti hubungan kausal

Jika ada sisi antara X dan Y, graf hanya mengatakan bahwa X dan Y dihubungkan menurut definisi graf tersebut. Sisi dapat berarti korelasi, kedekatan, komunikasi, konflik, ketergantungan model, atau relasi lain. Maknanya ditentukan oleh cara graf dibangun.

Kesalahan 2: Menyamakan “tidak ada sisi” dengan “tidak ada hubungan nyata”

Dalam data observasional, tidak ada sisi dapat berarti banyak hal:

- hubungan memang tidak ada;
- hubungan tidak diamati;
- hubungan tidak melewati ambang;
- hubungan dibuang saat

Document information

Bab 1: Objek Dasar Graf

Project	Teori Graf untuk Statistik
Document	Document 1.5
Author	Harizahayu
Verifier	Harizahayu
Downloaded	July 08, 2026 10:33 KST
Status	Verified
Document link	https://theorytrace.com/projects/teori-graf-untuk-statistik/documents/bab-1-objek-dasar-graf/