Butlletí de la Societat Catalana de Matemàtiques
Vol. 35, núm. 1, 2020. Pàg. 57–79. DOI: 10.2436/20.2002.01.93

Límits de grafs

Oriol Serra

Resum: La teoria dels límits de grafs es va desenvolupar la primera dècada del segle xxi amb l'objectiu de donar un model matemàtic de les grans xarxes que han aparegut amb l'eclosió d'Internet i que també són presents a la biologia, la física, la medicina o les ciències socials. La creació d'aquesta teoria ha estat liderada pel matemàtic hongarès László Lovász, que el febrer del 2019 va rebre el Premi Europeu de Ciència Hipàtia atorgat pel Barcelona Knowledge Hub de l'Academia Europaea. En aquest article es descriuen les nocions i resultats bàsics de la teoria de límits de grafs.

Paraules clau: grans xarxes, teoria de grafs, grafons, combinatòria extremal.

Classificació MSC2010: 05C35, 05C25, 05C80, 05C82, 90B15.

1 Introducció

En el moment d'escriure aquest article, la pàgina `https://www.internetlivestats.com/` indica que hi ha 1 705 226 930 pàgines web actives i 4 300 014 560 usuaris d'Internet. Xarxes d'aquesta mida no són infreqüents a la natura. S'estima que al cervell humà hi ha uns 100 miliards (100×10^9) de neurones. En deu grams de diamant hi ha uns 10^{23} àtoms cohesionats per una xarxa d'interaccions químiques, enriquides per una ínfima proporció d'impureses que caracteritzen el seu comportament. Tant des del punt de vista tecnològic com científic, el repte de tractar, entendre o gestionar xarxes amb un nombre molt gran de nodes s'ha convertit en un objectiu prioritari.

La teoria de grafs ofereix un model matemàtic natural per a les xarxes d'interconnexió. Un graf és un parell $G = (V, E)$ format per un conjunt V, els elements del qual s'anomenen *vèrtexs* o *nodes*, i un conjunt E de parells de vèrtexs, anomenats *arestes*. Un graf és un objecte matemàtic que formalitza les relacions binàries entre elements d'un conjunt. L'origen de la noció matemàtica de graf se situa en la solució que va donar Euler el 1736 a l'anomenat *problema dels ponts* de Königsberg, un problema de naturalesa topològica i combinatòria; vegeu, per exemple, [1]. El desenvolupament de la teoria de grafs es va produir sobretot al segle xx, impulsat d'una manera especial pel progrés de les ciències de la computació, però també per la seva versatilitat com a model natural en moltes àrees de la matemàtica, les ciències i la tecnologia.

El matemàtic hongarès László Lovász va ser un dels protagonistes del desenvolupament de la teoria de grafs en la segona meitat del segle XX, pel fet d'establir relacions profundes amb altres àrees de la matemàtica, com la topologia, la geometria, la probabilitat, l'àlgebra, l'anàlisi o l'optimització. Quan era professor a la Universitat de Yale, va acceptar una proposta per a incorporar-se al Theory Group of Microsoft Research a Redmond (Washington), on des del 1996 fins al 2006 va liderar el projecte de creació de la teoria matemàtica dels límits de grafs, esperonat pel repte de proporcionar una eina matemàtica per a l'anàlisi de grans xarxes. El resultat d'aquesta aventura està recollit en una extensa monografia, *Large Networks and Graph Limits* [11], publicada el 2012 per l'American Mathematical Society, i que està tenint un impacte enorme en l'estudi de les grans xarxes. El text estableix d'una banda el context teòric que ha esdevingut definitiu i, de l'altra, planteja diverses direccions de desenvolupament de la teoria que formen un projecte de recerca de gran abast.

Entre les múltiples distincions, premis i reconeixements que Lovász ha obtingut en la seva carrera professional com a matemàtic, trobem el Premi Europeu de Ciència Hipàtia, en la seva primera edició el 2019, instituït pel projecte conjunt de la Barcelona Knowledge Hub de l'Academia Europaea.

L'alcadessa de Barcelona Ada Colau en l'entrega del
Premi Hipàtia a László Lóvasz el 5 de febrer del 2019.

En el seu discurs d'acceptació, Lovász va expressar el seu convenciment que l'anàlisi de grans xarxes serà un dels motors del desenvolupament científic i tecnològic en les properes dècades. Entre les moltes aplicacions naturals de la teoria a les ciències de la computació, l'enginyeria, la biologia o la física, va aventurar que les eines d'anàlisi de grans xarxes havien de contribuir també a millorar el coneixement del món que ens envolta, vist des de l'òptica d'una immensa xarxa d'interaccions.

Aquest article vol oferir una introducció a la teoria de límits de grafs, esbossant les intuïcions que hi ha darrere l'elecció de les seves nocions bàsiques i descrivint els resultats més bàsics de convergència de grafs.

Per tal de presentar una visió global, hem optat per descriure la línia argumental que porta a les nocions bàsiques i hem obviat algunes de les demostracions que, o bé tenen una naturalesa molt tècnica, o bé tenen una complexitat que excedeix el format d'un article d'exposició com aquest. Dit això, hem volgut recollir tots els elements que condueixen a la definició de *successions convergents de grafs* i els seus límits, i presentar les proves d'alguns dels teoremes clau en aquesta teoria.

La teoria va molt més enllà del que s'exposa en aquest article. Les prop de cinc-centes pàgines de la monografia de Lovász [11], escrites amb la seva usual claredat i transparència, donen una perspectiva molt més àmplia del tema i són una referència excel·lent per a qualsevol lector interessat en el tema. Hi ha també exposicions més succintes que poden resultar útils, com les notes de Dan Král' [9] o l'article expositiu d'Andrzej Grzesik i Dan Král' [5].

Per situar el context, cal dir que hi ha dues teories generals de límits de grafs: la dels grafs anomenats *densos* i la dels grafs *dispersos*. Aquí ens centrarem només en la primera classe, grafs en els quals el nombre d'arestes és una proporció positiva del nombre total possible d'arestes. Si G és un graf amb $n = |V|$ vèrtexs, aleshores el nombre de possibles parells de vèrtexs és $\binom{n}{2} \sim n^2/2$. Els grafs densos tenen cn^2 arestes per a alguna constant $0 < c < 1/2$. Com que l'objectiu és tractar límits, aquestes afirmacions s'han d'entendre des del punt de vista asimptòtic. Els grafs dispersos, en canvi, tenen una quantitat subquadràtica d'arestes. No totes les grans xarxes es corresponen amb grafs densos, i alguns dels casos més interessants justament no ho són. Tot i així, la teoria de límits de grafs densos il·lustra d'una manera efectiva els problemes que es plantegen i el tipus de resultats que es persegueixen.

D'altra banda, els grafs són només un dels objectes combinatoris, potser el més simple, per al qual s'ha desenvolupat la teoria. L'objectiu general és establir una noció de límits d'estructures combinatòries, que inclouen permutacions, hipergrafs o complexos simplicials, entre d'altres. Fins i tot en l'àmbit dels grafs hi ha variants, com ara grafs etiquetats, grafs acolorits, grafs dirigits o grafs signats, que tenen un interès especial en les aplicacions. La teoria dels límits d'estructures combinatòries ha estat estesa a cadascun dels exemples mencionats prenent com a base la teoria dels límits de grafs.

En l'article es fa servir terminologia estàndard de teoria de grafs. El text de Reinhard Diestel [2] n'és una referència excel·lent. La majoria dels termes estan definits quan es fan servir per primera vegada. D'altres s'avenen a interpretacions naturals i confiem que el lector agrairà que s'obviïn algunes definicions formals. Per exemple, una *aresta* d'un graf és un parell de vèrtexs, i es diu també que l'aresta *uneix* els dos vèrtexs, que aquests li són *incidents* i que els dos vèrtexs són *adjacents*, totes elles denominacions que reflecteixen la representació gràfica usual d'un graf en el pla, en la qual els vèrtexs són punts i les arestes, arcs simples que els uneixen.

Un graf es pot definir sobre qualsevol conjunt, discret o no. En la teoria de grafs clàssica, hi ha molts problemes que tracten grafs amb conjunts de vèrtexs no numerables, per exemple, grafs geomètrics en espais euclidians.

A continuació, però, assumirem sempre que els grafs tenen un conjunt finit de vèrtexs. De fet, és usual imaginar que el terme G_n d'una successió de grafs $(G_n : n \in \mathbb{N})$ té n vèrtexs. La teoria de límits de grafs fa referència a successions de grafs en què les successions de nombres de vèrtexs dels termes de la successió no són fitades. El primer objectiu de la teoria és definir una bona noció de convergència d'una successió de grafs i identificar un objecte matemàtic adient per a la noció de límit d'una successió convergent.

2 Grafons

La noció de límit ha resultat fonamental en el desenvolupament de la matemàtica, ja que ha permès passar de l'enumeració a la mesura, del discret al continu. El conjunt de nombres reals és un concepte ideal construït a través de la noció de límit que, si bé té un referent difós a la natura i és sobretot una construcció intel·lectual, ha resultat d'una utilitat extraordinària. En la història de la matemàtica, la noció de límit ha estat estesa a objectes més complexos que els conjunts de nombres, com ara funcions, espais, estructures algebraiques, etc. El bagatge analític proporciona les eines necessàries per a saber entendre i manipular límits d'objectes matemàtics complexos. En el cas dels límits d'estructures combinatòries la novetat està en el fet que, tot i els precedents en el món de les estructures aleatòries discretes, no s'havia plantejat explícitament quins són els objectes límit adequats per a aquestes estructures.

Un dels objectius de la teoria de límits de grafs és capturar en un objecte límit algunes de les propietats raonables que pot interessar estudiar en l'anàlisi de grans xarxes: qüestions relatives a la connectivitat, l'existència de subestructures, la distribució de les arestes, particions de nodes amb graf quocient estructurat, etc. Per exemple, no té gaire sentit preguntar-se per la paritat del nombre de nodes o l'existència d'un nombre específic de triangles.

En el cas dels grafs, que descriuen relacions binàries en un conjunt, la definició de *límit* que resulta natural és la següent:

Definició 1 (Grafó). Un *grafó* és una funció mesurable (entès sempre respecte de la mesura de Lebesgue)

$$W\colon [0,1]\times[0,1] \to [0,1],$$

que és simètrica en les seves variables: $W(x,y) = W(y,x)$ per a tot $(x,y) \in [0,1]\times[0,1]$.

El nom *grafó* prové de la contracció de *graf* i *funció*. Denotem per $\mathcal{W}_0$ l'espai d'aquestes funcions, és a dir, l'espai de funcions simètriques de $L^\infty([0,1]\times[0,1])$ que prenen valors a $[0,1]$.

Donat un graf G, es pot definir un grafó associat a G de la manera següent. Si G té un conjunt de vèrtexs $\{1,2,\dots,n\}$, considerem una partició de $[0,1]$ en intervals $I_1 = [0,1/n), I_2 = [1/n, 2/n), \dots, I_n = [(n-1)/n, 1]$ de la mateixa mida. Denotem per W_G el grafó

$$W_G(x,y) = \begin{cases} 1 & (x,y) \in I_i \times I_j,\ \{i,j\} \in E(G),\\ 0 & (x,y) \in I_i \times I_j,\ \{i,j\} \notin E(G).\end{cases}$$

De fet, qualsevol partició de $[0,1]$ en n intervals, o n conjunts mesurables, serviria per a fer aquesta associació, però la nostra notació per al grafó W_G estarà relacionada amb una equipartició de l'interval. A la figura 1 s'il·lustra aquesta associació per al cas del graf bipartit complet $K_{5,5}$, en què el conjunt de vèrtexs està partit en dos subconjunts de cinc vèrtexs cadascun i les seves arestes són tots els parells amb un vèrtex a cada part.

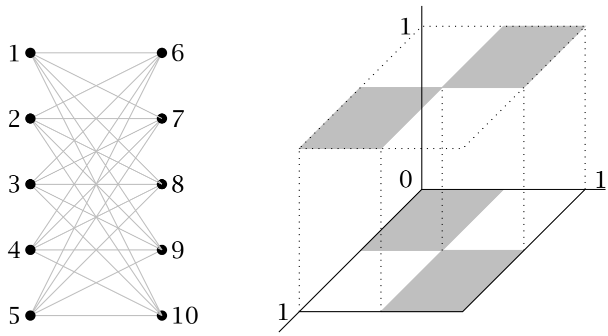
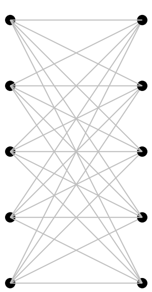


Figura 1: El graf bipartit complet $K_{5,5}$ i el seu grafó associat.

Així doncs, a cada graf se li pot associar un grafó de manera natural, que pren valors a $\{0,1\}$. Veurem que en la definició de *convergència de grafs* no n'hi ha prou amb els grafons que prenen valors a $\{0,1\}$. Totes les funcions de $\mathcal{W}_0$ són necessàries per a completar l'espai dels grafs.

Una manera gràfica de relacionar grafs i grafons s'obté a través de la matriu d'adjacència del graf. Si $G = (V, E)$ és un graf amb un conjunt de vèrtexs $V = \{1, 2, \ldots, n\}$, la matriu d'adjacència de G és una matriu quadrada $A = A(G)$ d'ordre n amb entrades a $\{0,1\}$, on $A_{i,j}$ val 1 si $\{i, j\}$ és una aresta del graf i $A_{i,j} = 0$ altrament. La figura 2 il·lustra el graf complet bipartit $K_{5,5}$ i la seva matriu d'adjacència.



$$\begin{pmatrix}
0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1\\
0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1\\
0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1\\
0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1\\
0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1\\
1 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0\\
1 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0\\
1 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0\\
1 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0\\
1 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0
\end{pmatrix}$$

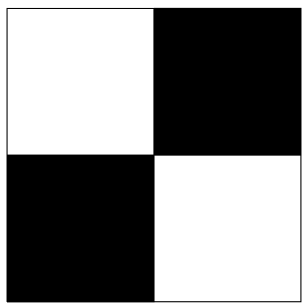


Figura 2: El graf bipartit complet $K_{5,5}$, la seva matriu d'adjacència i el grafó complet bipartit.

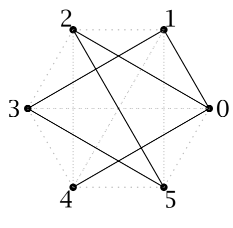
Més endavant veurem que la successió $\{K_{r,r},\, r \in \mathbb{N}\}$ de grafs bipartits complets és convergent i té per límit el grafó bipartit complet $W_{K_{1,1}}$ definit com

$$W(x,y) = \begin{cases} 1 & (x,y) \in [0,1/2) \times [1/2,1], \\ 0 & \text{altrament.} \end{cases}$$

S'obté una representació gràfica d'un grafó dibuixant al quadrat $[0,1]\times[0,1]$ les imatges com a punts del blanc (0) al negre (1) passant per diferents intensitats de grisos per a valors a $(0,1)$. D'aquesta manera es pot interpretar el grafó com un límit de les matrius d'adjacència quan l'ordre n de la matriu tendeix a infinit. Amb aquesta representació, el grafó bipartit complet està representat a la figura 2 (observem que, respectant la notació matricial, el sentit creixent de l'ordenada és de dalt a baix).

Una altra manera de relacionar grafs i grafons ve donada pels grafs aleatoris. L'anomenat *graf aleatori* d'Erdős-Rényi d'ordre n i probabilitat p, $G(n,p)$, s'obté posant cada aresta $\{i,j\}$ independentment amb probabilitat p en el conjunt de vèrtexs $\{1,2,\dots,n\}$. Per a $p = 1/2$ és fàcil comprovar que l'objecte aleatori que resulta correspon a escollir aleatòriament amb la distribució uniforme un graf entre tots els possibles grafs de n vèrtexs.

A cada grafó W se li pot associar un graf aleatori G_W de n vèrtexs de la manera següent. Un W-graf aleatori d'ordre n és el graf G_W que té per conjunt de vèrtexs n punts $x_1,\dots,x_n$ escollits independentment amb la distribució uniforme a $[0,1]$, i posant cada aresta $\{x_i,x_j\}$ a G_W amb probabilitat $W(x_i,x_j)$ independentment. Quan W és un grafó constant, $W \equiv p$, aleshores G_W és el graf aleatori $G(n,p)$. Veurem que la successió $\{G(n,1/2) : n \in \mathbb{N}\}$ de grafs aleatoris amb $p = 1/2$ convergeix al grafó $W \equiv 1/2$. La figura 3 il·lustra una possible realització del graf aleatori $G(6,1/2)$ i el grafó $W \equiv 1/2$.



$$\begin{pmatrix} 0 & 1 & 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 \end{pmatrix}$$

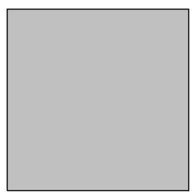


Figura 3: Una realització del graf aleatori $G(6,1/2)$, la seva matriu d'adjacència i el grafó $W \equiv 1/2$.

A continuació es donarà una noció natural de *convergència de successions de grafs* i es veurà que els grafons són els objectes límit naturals per a aquesta noció.

3 Homomorfismes de grafs

La definició dels grafons com a objectes naturals de límits de successions de grafs té diverses justificacions. Una de les més il·lustratives prové de la noció d'homomorfisme de grafs. Donats dos grafs G i H, una aplicació $f\colon V(G) \to V(H)$ és un *homomorfisme de grafs* si respecta les adjacències: per a cada aresta $\{x, y\}$ de G, $\{f(x), f(y)\}$ és una aresta de H. Aquesta és la noció natural de morfisme en la categoria dels grafs. Si hi ha un homomorfisme del graf G al graf H, escrivim $G \to H$. Observem que un homomorfisme $f\colon G \to H$ pot enviar un parell $\{x, y\}$ que no és una aresta de G a un parell $\{f(x), f(y)\}$ que pot ser o no una aresta de H.

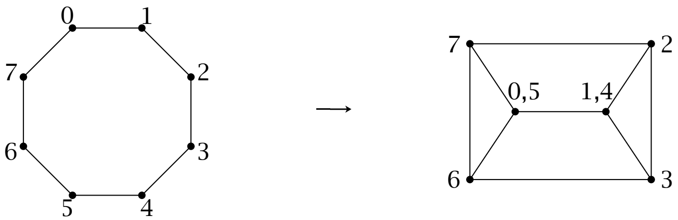


Figura 4: Un homomorfisme del cicle C_8 al graf de la dreta.

Si $f\colon G \to H$ és un homomorfisme bijectiu tal que el seu invers és també un homomorfisme, aleshores f és un *isomorfisme* i els grafs G i H són isomorfs, és a dir, indistingibles llevat d'una permutació dels vèrtexs.

Els homomorfismes permeten descriure moltes de les propietats d'un graf. Per exemple, si K_k denota el graf complet amb k vèrtexs (tots els parells de $\{1, 2, \ldots, k\}$ són arestes de K_k), aleshores $K_k \to G$ indica que el graf G conté una *clica d'ordre k* (un subgraf de G isomorf a K_k). D'altra banda, $G \to K_k$ indica que G admet una k-coloració pròpia, és a dir, una partició dels vèrtexs de G en k parts de manera que cada part indueix un subgraf sense arestes. En altres paraules, si $f\colon G \to K_k$ és un homomorfisme i acolorim cada vèrtex $x \in V(G)$ amb el «color» $f(x)$, cada aresta de G és incident amb vèrtexs de colors diferents (no hi ha arestes monocromàtiques). Un graf G que és 2-acolorible es diu que és un graf bipartit: el conjunt de vèrtexs de G admet una partició en dues parts de manera que les arestes de G tenen un sol vèrtex a cada part. No és difícil comprovar que si G és un graf bipartit connex (cada parell de vèrtexs es pot unir per una successió de vèrtexs cadascun adjacent al següent), aleshores la bipartició és única.

La complexitat algorísmica del problema de decidir, donats dos grafs G i H, si $G \to H$, reflecteix la complexitat dels homomorfismes. Si el graf de sortida G és fixat, decidir per a cada H si $G \to H$ és un problema que es pot resoldre en temps que és polinòmic en la mida de l'entrada H del problema. Més específicament, el problema és polinòmic en el nombre $n = |V(H)|$ de vèrtexs de H: només cal inspeccionar cadascuna de les aplicacions de $V(G) \to V(H)$, de les quals n'hi ha $n^{|V(G)|}$ (una quantitat polinòmica en n) i comprovar si és un homomorfisme. Un teorema cèlebre de Pavol Hell i Jaroslav Nešetřil ([6])

estableix que, si fixem el graf d'arribada H, el problema de decidir per a cada G si $G \to H$ és polinòmic en la mida $n = |V(G)|$ de l'entrada del problema només si H és bipartit, i és en canvi NP-complet si H no és bipartit. Per als lectors no familiaritzats amb la complexitat algorísmica, n'hi ha prou de saber que la classe de problemes NP-complets es considera a la cúspide d'una jerarquia de dificultat en la resolució d'una àmplia classe de problemes de manera algorísmica. Segons el teorema, doncs, per a H fix, decidir si $G \to H$ és o bé senzill o bé molt difícil. Per això el resultat és conegut com la dicotomia dels homomorfismes. La monografia [7] dels mateixos autors és una referència excel·lent sobre homomorfismes de grafs.

Denotem per $\mathrm{Hom}(G,H)$ el conjunt de tots els homomorfismes de G en H i amb $\hom(G,H)$ el seu nombre. Ocasionalment serà útil considerar només homomorfismes injectius, el nombre dels quals es denota per $\mathrm{inj}(G,H)$. També serà útil el nombre dels homomorfismes injectius que respecten la no adjacència, és a dir, homomorfismes injectius $f\colon G \to H$ tals que, per a cada parell de vèrtexs x, y de G, la imatge $\{f(x), f(y)\}$ és una aresta de H si, i només si, $\{x,y\}$ és una aresta de G. En aquest cas, $f(G)$ és un subgraf *induït* de H, és a dir, hi ha un subconjunt de vèrtexs $U \subset V(H)$ tal que el subgraf $H[U]$ de H format pels vèrtexs de U és isomorf a G. Diem aleshores que f és una *immersió* de G en H i denotem per $\mathrm{ind}(G,H)$ el nombre d'immersions de G en H.

Les tres quantitats, $\hom(G,H)$, $\mathrm{inj}(G,H)$ i $\mathrm{ind}(G,H)$, estan relacionades entre elles. Per exemple,

$$\mathrm{inj}(G,H) = \sum_{G' \supseteq G} \mathrm{ind}(G',H),$$

on el sumatori s'estén a tots els grafs G' que contenen G amb el mateix conjunt de vèrtexs, i

$$\hom(G,H) = \sum_{P} \mathrm{inj}(G/P,H),$$

on el sumatori s'estén a totes les particions P del conjunt de vèrtexs de G i G/P denota el graf quocient, que té per vèrtexs els elements de P i dues parts són adjacents si hi ha una aresta de G que uneix un vèrtex de cada part. Les relacions inverses es poden obtenir fent servir la fórmula d'inversió de Möbius.

Per a un graf G, el conjunt de nombres $\hom(\cdot,G) = \{\hom(H,G) : H \in \mathcal{G}\}$, on $\mathcal{G}$ és la classe de tots els grafs, s'anomena *perfil de G per l'esquerra*. De manera anàloga, $\hom(G,\cdot) = \{\hom(G,H) : H \in \mathcal{G}\}$ és el *perfil de G per la dreta*. Un resultat rellevant, obtingut per Lovász fa més de cinquanta anys, és el següent.

Teorema 2 (Lovász [10]). *Els nombres* $\hom(\cdot,G)$ *determinen G llevat d'isomorfisme. El mateix és cert per als nombres* $\hom(G,\cdot)$.

Prova. Si G i G' tenen el mateix perfil per l'esquerra, se satisfà també que

$$\mathrm{inj}(H,G) = \mathrm{inj}(H,G'),$$

per a tots els grafs H. En particular,

$$\mathrm{inj}(G,G) = \mathrm{inj}(G,G') = \mathrm{inj}(G',G) = \mathrm{inj}(G',G'),$$

i tots quatre nombres són no nuls. Per tant, hi ha un homomorfisme injectiu de G en G' i un de G' en G. Així doncs, els dos grafs són isomorfs. Un argument semblant prova el cas del perfil per la dreta. □

Els nombres $\hom(G, H)$ permeten expressar moltes de les propietats d'un graf. De l'extensa llista de relacions entre paràmetres de grafs i nombres d'homomorfismes, en donem només alguns exemples senzills (les denominacions dels grafs implicats estan il·lustrades a la figura 5).

1. Denotem per S_k l'estrella de k vèrtexs. El conjunt d'arestes de S_k és $\{\{1, i\} : i = 2, \ldots, k-1\}$. El grau $d(x)$ d'un vèrtex x en un graf G és el nombre d'arestes de G que són incidents amb x. El nombre $\hom(S_k, G)$ dona els moments d'ordre $k-1$ dels graus:
$$\hom(S_k, G) = \sum_{x \in V(G)} (d(x))^{k-1}.$$
2. Denotem per C_k el cicle de k vèrtexs. El conjunt d'arestes de C_k és $\{\{i, i+1 \pmod k\} : i = 1, \ldots, k\}$. Aleshores,
$$\hom(C_k, G) = \mathrm{tr}(A(G)^k) = \sum_{i=1}^{n} \lambda_i^k,$$
on $\mathrm{tr}(A(G))$ denota la traça de la matriu d'adjacència de G i $\lambda_1, \ldots, \lambda_n$ són els valors propis de $A(G)$ (com que és real i simètrica, la matriu d'adjacència diagonalitza i té n valors propis reals).
No és difícil comprovar que $\lim_{k\to\infty} \hom(C_{2k}, G)^{1/2n}$ existeix i que el seu valor és el valor propi més gran de $A(G)$.
3. Si K_k denota el graf complet de k vèrtexs, $\hom(G, K_k)$ és el nombre de coloracions pròpies de G.
4. Si H_2 denota el graf complet K_2 amb un llaç a un dels dos vèrtexs, aleshores $\hom(G, H_2)$ és el nombre de conjunts estables de G, subconjunts $U \subset V(G)$ tals que el subgraf induït per U no té cap aresta.

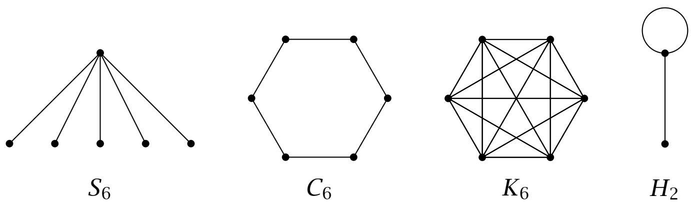


Figura 5: Exemples d'estrelles, cicles, grafs complets i H_2.

Aquests exemples il·lustren com els nombres $\hom(H, G)$ recullen informació estructural del graf G en relació amb el graf H. L'estudi dels límits de grafs, però, no pot ser sensible només a la noció de grafs isomorfs, la qual cosa donaria una eina poc útil, sinó que ha de comportar un compromís entre propietats locals i propietats globals que puguin ser reflectides en el límit.

4 Convergència de successions de grafs

El paràmetre que ens interessarà de cara a formular la noció de convergència de grafs és la *densitat* d'homomorfismes. Si G és un graf amb n vèrtexs i H és un graf amb k vèrtexs, la densitat d'homomorfismes de H en G és

$$t(H,G) = \frac{\hom(H,G)}{n^k}.$$

En altres paraules, $t(H,G)$ és la probabilitat que una aplicació aleatòria de H en G sigui un homomorfisme. De manera anàloga es defineixen les densitats d'homomorfismes injectius i d'immersions,

$$t_{\text{inj}}(H,G) = \frac{\text{inj}(H,G)}{(n)_k} \quad \text{i} \quad t_{\text{ind}}(H,G) = \frac{\text{ind}(H,G)}{(n)_k},$$

on $(n)_k = n(n-1)\cdots(n-k+1)$ és el nombre d'aplicacions injectives de $V(H)$ en $V(G)$. En particular, $t_{\text{ind}}(G,H)$ és la probabilitat que una aplicació aleatòria injectiva de $V(H)$ en $V(G)$ tingui com a imatge un subgraf induït de G isomorf a H, és a dir, una còpia de H en G.

De la mateixa manera que els conjunts de nombres $\hom(\cdot,G)$, $\hom_{\text{inj}}(\cdot,G)$ i $\hom_{\text{ind}}(\cdot,G)$ estan relacionats entre ells, es pot obtenir cadascun dels conjunts $t(\cdot,G)$, $t_{\text{ind}}(\cdot,G)$ i $t_{\text{ind}}(\cdot,G)$ de qualsevol dels altres dos per relacions lineals entre els seus membres.

A diferència dels nombres $\hom(\cdot,G)$, els nombres de densitats $t(\cdot,G)$ no determinen el graf G llevat d'isomorfisme. El motiu és que, si se substitueix cada vèrtex de G per r còpies, mantenint les adjacències, aleshores el graf resultant té el mateix conjunt de densitats. Més precisament, donat un graf G i un enter positiu r, denotem per $G(r)$ el graf obtingut de G substituint cada vèrtex $x \in V(G)$ per r vèrtexs $x_1,\ldots,x_r$, i x_i, y_j són adjacents a $G(r)$ si, i només si, x, y són adjacents a G. Es diu que $G(r)$ és un graf *expandit* de G. De la definició de *densitat d'homomorfismes* es pot comprovar fàcilment que

$$t(\cdot,G) = t(\cdot,G(r)).$$

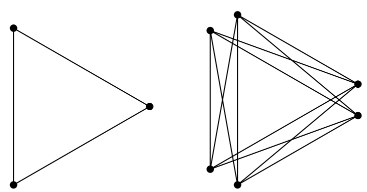


Figura 6: El graf K_3 i el graf expandit $K_3(2)$.

No és difícil verificar que el recíproc també és cert: si $t(\cdot,G) = t(\cdot,G')$, aleshores G és isomorf a $G'(r)$ per a algun r o viceversa, G' és isomorf a $G(r)$ per a algun r.

Amb aquestes definicions es pot formular la noció de *convergència d'una successió de grafs*.

Definició 3. Una successió de grafs $(G_n : n \in \mathbb{N})$ amb $\lim_{n\to\infty} |V(G_n)| = \infty$ és convergent si, per a cada graf H, la successió

$$(t(H, G_n) : n \in \mathbb{N})$$

és convergent.

Recordem que els nombres $t(\cdot, G)$ de densitats d'homomorfismes i $t_{\mathrm{ind}}(\cdot, G)$ de densitats d'immersions estan relacionats de manera que es poden obtenir l'un de l'altre. Per tant, una successió de grafs és convergent si, i només si, les successions $t_{\mathrm{ind}}(H, G)$ són convergents per a cada H. Aquesta és la intuïció que hi ha darrere la definició de *convergència*. Es mesura l'estructura d'un graf en termes dels estadístics dels subgrafs que conté. Si dos grafs tenen la mateixa densitat d'immersions d'un graf donat H, és a dir, el mateix nombre de còpies de H, aleshores els grafs són semblants en relació amb H. Si aquest és el cas per a tots els grafs H, aleshores els dos grafs són indistingibles.

Per exemple, resulta immediat comprovar que la successió de grafs complets $(K_n : n \in \mathbb{N})$ és convergent. Per a cada graf H, cada aplicació injectiva $f : V(H) \to V(K_n)$ és un homomorfisme, de manera que $t_{\mathrm{inj}}(H, K_n) = 1$ per a tot n i la successió $(t_{\mathrm{inj}}(H, K_n) : n \in \mathbb{N})$ és convergent. Un altre exemple és el dels grafs aleatoris $G(n, p)$. Per a $p \in (0, 1)$ i un graf H fixats, la probabilitat que una aplicació aleatòria injectiva $f : V(H) \to V(G(n, p))$ sigui un homomorfisme és igual a $p^{|E(H)|}$, per a $n \geq |V(H)|$, i la successió $t_{\mathrm{inj}}(H, G(n, p))$ és convergent per a cada H. Un tercer exemple és la successió de grafs bipartits complets $(K_{n,n} : n \in \mathbb{N})$. En aquest cas, si H no és bipartit, aleshores $t(H, K_{n,n}) = 0$ per a tot n (un graf bipartit es caracteritza perquè conté un cicle imparell i la imatge per un homomorfisme d'un cicle imparell és un cicle imparell), mentre que si H és bipartit amb bipartició $V(H) = A \cup B$, aleshores una aplicació $f : V(H) \to V(K_{n,n})$ és un homomorfisme si, i només si, envia A i B a parts diferents de $V(K_{n,n})$, de manera que $t_{\mathrm{inj}}(H, K_{n,n}) = 1/2^{|A|+|B|-1}$. La bipartició d'un graf bipartit és única només si el graf és connex, però una modificació de l'argument anterior fa veure que, per a qualsevol graf bipartit H, connex o no, $t(H, K_{n,n})$ és una constant que depèn només de H per a tot n prou gran. Per tant, la successió $(K_{n,n} : n \in \mathbb{N})$ és convergent.

Observem que la definició de *convergència* que hem donat és rellevant només per a grafs densos. El motiu és que, en la normalització que s'ha fet servir en la definició de la densitat d'homomorfismes, si el nombre d'arestes dels termes de la successió és $o(n^2)$, aleshores $t(H, G_n)$ convergeix trivialment a zero per a cada H fixat. En altres paraules, si G_n és una successió de grafs dispersos, aleshores és trivialment convergent.

Tot i que una noció de convergència adequada és suficient per a definir una teoria de límits, perquè una definició com aquesta sigui satisfactòria cal aclarir algunes qüestions. En primer lloc, quin és el significat d'aquesta convergència en termes d'una mètrica adequada. Per tal que una successió de grafs sigui convergent, cal que els termes de la successió siguin semblants quan n és gran i que la seva distància respecte d'alguna mètrica tendeixi a zero. Quina és aquesta mètrica i què representa en termes de comparar dos grafs grans? En segon lloc, si una successió és convergent s'ha de poder identificar el límit i aquest ha de ser únic. Ja hem avançat que l'objecte límit és un grafó. Novament, cal definir una mètrica per la qual els termes d'una successió de grafs convergent estan a una distància cada vegada més petita d'un grafó. Finalment, si l'espai dels grafons ha de ser la compleció de l'espai dels grafs, cada grafó hauria de ser el límit d'una successió convergent de grafs. A continuació veurem com es dona resposta a aquestes preguntes.

5 La distància de tall

L'objectiu és definir una mètrica en l'espai $\mathcal{G}$ dels grafs que mesuri la diferència estructural entre dos grafs. Aquesta mètrica es defineix primer per a grafs sobre el mateix conjunt de vèrtexs i s'estén després a parells arbitraris de grafs.

Una manera natural de comparar dos grafs és mesurar el nombre d'arestes que cal modificar per a transformar l'un en l'altre. Convenientment normalitzada, aquesta mesura dona lloc al que s'anomena *distància d'edició*. Donats dos grafs G, G' amb el mateix conjunt de vèrtexs, la distància d'edició entre G i G' és

$$d_1(G, G') = \frac{|E(G)\Delta E(G')|}{n^2},$$

on Δ denota la diferència simètrica de conjunts i n és el nombre de vèrtexs.

Una mesura més sensible a l'estructura dels grafs és el que s'anomena *distància de tall*. Per a dos subconjunts de vèrtexs S, T denotem per $e_G(S,T) = \sum_{i\in S,\, j\in T} a_{ij}$, on a_{ij} és l'entrada de la matriu d'adjacència $A(G)$, el nombre d'arestes que uneixen vèrtexs de S amb vèrtexs de T en el graf G. Comencem per definir el que s'anomena *distància de tall etiquetada*. Donats dos grafs G, G' amb el mateix conjunt de vèrtexs, la distància de tall etiquetada entre G i G' és

$$\delta_\square(G, G') = \max_{S,T\subset V(G)} \frac{|e_G(S,T) - e_{G'}(S,T)|}{n^2},$$

on n és el nombre de vèrtexs.

No és difícil comprovar que d_1 i $\delta_\square$ són efectivament distàncies en el conjunt de grafs sobre el mateix conjunt de vèrtexs i que

$$\delta_\square(G, G') \le d_1(G, G'),$$

tot i que les dues distàncies poden ser ben diferents.

Des del punt de vista estructural, la distància entre dos grafs isomorfs hauria de ser zero. Resulta natural doncs definir, per a dos grafs G i G' amb el mateix nombre de vèrtexs, la menor distància entre grafs isomorfs a l'un i a l'altre:

$$\widehat{\delta_\square}(G, G') = \min_{H\cong G,\, H'\cong G'} \delta_\square(H, H').$$

No és difícil comprovar que $\widehat{\delta_\square}$ és una distància en el conjunt de classes de grafs isomorfs. És clar que $\widehat{\delta_\square}(G,G') \leq \delta_\square(G,G')$.

Finalment, si G té n vèrtexs i G' té n' vèrtexs, aleshores els grafs expandits $G(n')$ i $G'(n)$ tenen el mateix nombre de vèrtexs, i els grafs G i $G(n')$ són estructuralment semblants. En particular, $t(\cdot,G) = t(\cdot,G(n'))$. D'aquesta manera s'arriba a la definició de la *distància de tall* (*cut-distance* en anglès).

DEFINICIÓ 4 (DISTÀNCIA DE TALL). La distància de tall entre dos grafs G i G' és

$$d_\square(G,G') = \lim_{k\to\infty} \widehat{\delta_\square}(G(kn'),G'(kn)),$$

on n és el nombre de vèrtexs de G i n' és el nombre de vèrtexs de G'.

La denominació *distància de tall* prové del fet que la distància es mesura en termes del nombre d'arestes que uneixen dos subconjunts S i T de vèrtexs, és a dir, del nombre d'arestes que s'han de tallar per a separar aquests dos conjunts en el graf.

Una vegada més, es pot comprovar que $d_\square(G,G') = 0$ si, i només si, hi ha un graf H i dos nombres k, k' tals que $G \cong H(k)$ i $G' \cong H(k')$, de manera que $d_\square$ és una distància en el conjunt de classes de grafs equivalents per a aquesta noció d'isomorfia.

Equipats amb aquesta mètrica, es pot formular el teorema que dona resposta a la primera pregunta sobre successions convergents.

TEOREMA 5. *Sigui* $(G_n : n \in \mathbb{N})$ *una successió de grafs convergent. Aleshores la successió és de Cauchy respecte de la distància de tall, és a dir, per a tot enter k,*

$$\lim_{n\to\infty} d_\square(G_n,G_{n+k}) = 0.$$

Com passa en altres processos de compleció, la prova del teorema anterior és més abordable i transparent quan es pot fer referència al límit de la successió. Així doncs, abans de discutir la prova del teorema, discutirem com s'estén la mètrica de la distància de tall i les densitats d'homomorfismes als grafons.

6 Distància de tall entre grafons

La distància de tall es pot estendre de manera natural al conjunt de grafons. De fet, en el context dels grafons, les definicions són més transparents i els resultats més senzills de provar.

Donats dos grafons W, W', la distància de tall etiquetada entre W i W' és

$$\delta_\square(W,W') = \sup_{S,T\subset[0,1]} \left| \int_{S\times T} (W(x,y) - W'(x,y))\,dx\,dy \right|.$$

Com s'ha fet en el cas dels grafs, la distància hauria de reflectir una propietat estructural i ser, per tant, invariant per simetries. Anomenem $S_{[0,1]}$ el grup de funcions invertibles $\phi\colon [0,1] \to [0,1]$ que preserven la mesura. Per a un grafó W i $\phi \in S_{[0,1]}$, denotem per W^ϕ el grafó definit com

$$W^\phi(x,y) = W(\phi(x),\phi(y)).$$

Dos grafons W, W' són isomorfs si $W' = W^\phi$ per a alguna $\phi \in S_{[0,1]}$. Per exemple, els grafons de la figura 7 són isomorfs a través de la transformació

$$\phi(x) = \begin{cases} 1-x & x \in (1/4, 3/4), \\ x & \text{altrament.} \end{cases}$$

De manera informal, tots dos són límits dels grafs complets bipartits, com s'ha suggerit a la figura 2.

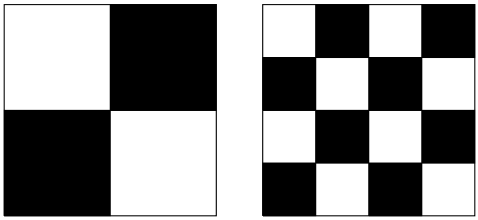


Figura 7: Dos grafons isomorfs.

Definició 6 (Distància de tall). Donats dos grafons W, W', la distància de tall entre W i W' és

$$d_\square(W,W') = \inf_{\phi,\phi' \in S_{[0,1]}} \delta_\square(W^\phi, (W')^{\phi'}).$$

Es pot provar que l'ínfim en la definició anterior és assolit per algun parell $\phi, \phi' \in S_{[0,1]}$. Amb una mica més de feina es pot provar també que dos grafons estan a distància zero si, i només si, són isomorfs. Així, $d_\square$ és una distància en el conjunt de classes d'isomorfia de grafons. Un resultat important és que l'espai mètric que en resulta és compacte.

Teorema 7 (Lovász i Szegedy [12]). *L'espai mètric dels grafons amb la distància de tall $d_\square$ és compacte.*

7 Densitat de grafs en grafons

El darrer ingredient per a relacionar grafs i grafons és la densitat d'immersions d'un graf H en un grafó W. La idea és la mateixa que en la definició de la densitat d'immersions de H en un graf G: la densitat d'immersions mesura el nombre de còpies de H en W.

El que cal és definir què s'entén per una còpia de H en W. Recordem que, donat un grafó W, un W-graf aleatori d'ordre n és el graf que s'obté escollint n punts $x_1, \ldots, x_n$ amb la distribució uniforme a $[0,1]$ i construint el graf G_W amb n vèrtexs posant cada aresta $\{x_i, x_j\}$ a G amb probabilitat $W(x_i, x_j)$ independentment. La mesura del nombre de còpies de H en W és la probabilitat que el graf G_W d'ordre $n = |V(H)|$ sigui isomorf a H; més precisament, tenim la definició següent.

Definició 8 (Densitat). Donat un graf H amb un conjunt de vèrtexs $\{1, 2, \ldots, n\}$ i un grafó W, la densitat d'immersions de H en W és

$$t_{\mathrm{ind}}(H, W) = \frac{n!}{|\mathrm{Aut}(H)|} \int_{[0,1]^n} \prod_{ij \in E(H)} W(x_i, x_j) \prod_{ij \notin E(H)} (1 - W(x_i, x_j))\, dx_1 \cdots dx_n.$$

El terme $n!/|\mathrm{Aut}(H)|$ en la definició anterior té en compte les possibles $n!$ immersions en un conjunt donat de n punts, normalitzat pel nombre $|\mathrm{Aut}(H)|$ d'automorfismes de H, que donen lloc a la mateixa còpia de H. La integral mesura la probabilitat que les arestes de H siguin presents en la immersió i que les no arestes de H no hi figurin.

De la definició de *grafons isomorfs* es conclou fàcilment que si dos grafons W, W' són isomorfs, aleshores $t_{\mathrm{ind}}(H, W) = t_{\mathrm{ind}}(H, W')$, ja que les funcions invertibles que preserven la mesura deixen invariant la integral.

De forma anàloga a com hem definit una successió convergent de grafs, es defineix la convergència d'una successió de grafs a un grafó.

Definició 9 (Successió convergent). Una successió $(G_n : n \in \mathbb{N})$ convergeix a un grafó W si la successió

$$(t_{ind}(G_n, W) : n \in \mathbb{N})$$

és convergent.

Veurem que la definició anterior és equivalent a la definició 3 de successions convergents de grafs, és a dir, que una successió de grafs és convergent segons la definició 3 si, i només si, hi ha un grafó W tal que la successió hi convergeix. Abans, però, veiem que cada grafó és el límit d'una successió convergent de grafs. Aquesta successió és precisament la dels W-grafs aleatoris d'ordre n associats al grafó W. La idea de la prova consisteix a anar revelant un a un els vèrtexs d'un W-graf aleatori per relacionar les quantitats $t(H, G_n)$ i $t(H, W)$ per a cada graf fixat H, on G_n és un W-graf aleatori d'ordre n. La prova fa servir una desigualtat de concentració per a martingales coneguda com la desigualtat d'Azuma-Hoeffding i el lema de Borel-Cantelli, que recordem a continuació.

Una successió $(X_n : n \in \mathbb{N})$ de variables aleatòries amb esperança finita sobre un espai de probabilitat és una martingala si, per a cada $n > 1$, l'esperança de X_n condicionada a $X_1, \ldots, X_{n-1}$ és X_{n-1}, és a dir,

$$\mathbb{E}(X_n | X_1, \ldots, X_{n-1}) = X_{n-1}.$$

Un cèlebre resultat de concentració en martingales amb valors de salt fitats és la desigualtat d'Azuma-Hoeffding (vegeu, per exemple, la monografia de Colin McDiarmid [14] sobre desigualtats de concentració).

TEOREMA 10 (AZUMA-HOEFFDING). *Sigui* $(X_n : n \in \mathbb{N})$ *una martingala amb* $\mathbb{E}(X_n) = x_0$ *i sigui* c_n *una successió de nombres reals tals que*

$$\Pr(|X_n - X_{n-1}| \leq c_n) = 1.$$

Aleshores, per a cada nombre real positiu t *i cada* $n \geq 1$

$$\Pr(|X_n - x_0| \geq t) \leq e^{-\frac{t^2}{2\sum_{i=1}^{n} c_i^2}}.$$

El lema de Borel-Cantelli és un resultat bàsic de teoria de probabilitat (i de teoria de la mesura) que es pot enunciar de la manera següent.

LEMA 11 (BOREL-CANTELLI). *Sigui* $(A_n : n \in \mathbb{N})$ *una successió d'esdeveniments en un espai de probabilitat. Si*

$$\sum_{n \geq 1} \Pr(A_n) < \infty,$$

aleshores la probabilitat que se satisfaci una quantitat infinita d'esdeveniments de la successió és zero.

A partir d'aquests resultats, es pot provar el teorema següent.

TEOREMA 12. *Sigui* W *un grafó. La successió* G_n *de grafs on cada* G_n *és un* W*-graf aleatori d'ordre* n *és convergent i el seu límit és* W *amb probabilitat* 1*.*

PROVA. Fixem un graf H amb $h = |V(H)|$ vèrtexs i $n > h$. Observem primer que, per la definició de G_n com a W-grafó i la definició de $t(H, W)$, la probabilitat que un subconjunt particular de h vèrtexs de G_n indueixi una còpia de H és $t(H, W)$.

Definim ara una successió $X_0, X_1, \ldots, X_n$ de variables aleatòries, on X_i denota el nombre esperat de còpies de H en G_n quan s'han fixat i vèrtexs de G_n i les arestes del graf induït per aquests i vèrtexs en la construcció del W-graf aleatori G_n. Per la linealitat de l'esperança, $X_0 = \binom{n}{h} t(H, W)$ i $X_n = \binom{n}{h} t(H, G_n)$.

Observem que $\mathbb{E}(X_i | X_{i-1}) = \mathbb{E}(X_{i-1})$, de manera que la successió de variables aleatòries és una martingala. D'altra banda, $|X_i - X_{i-1}| \leq n^{h-1}$ per a cada $i = 1, \ldots, n$, fitant generosament pel nombre possible d'aplicacions de $[h]$ en $[n]$ quan s'ha fixat almenys un punt. Fent servir la desigualtat d'Azuma-Hoeffding, per a cada nombre real positiu t,

$$\Pr(|X_n - X_0| \geq t) \leq 2e^{-\frac{t^2}{2n^{2h-1}}}.$$

Per a cada $\epsilon > 0$, posant $t = \epsilon n^h$ a la desigualtat anterior, s'obté

$$\Pr(|X_n - X_0| \geq \epsilon n^h) \leq 2e^{-\epsilon^2 n/2}.$$

Utilitzant la fitació $\binom{n}{h} \geq n^h/h^h$, obtenim finalment

$$\Pr(|t(H, G_n) - t(H, W)| \geq h^h \epsilon) \leq 2e^{-\epsilon^2 n/2}.$$

Pel lema de Borel-Cantelli, la successió $t(H, G_n)$ és convergent i el seu límit és $t(H, W)$ amb probabilitat 1. □

8 El lema de regularitat

Per a concloure l'equivalència entre les nocions de convergència de successions de grafs i la convergència a un grafó, cal una darrera peça que enllaça amb la noció de distància de tall. Aquesta peça és un resultat celebrat de Szemerédi que ha tingut un impacte important en el desenvolupament de la combinatòria i la informàtica teòrica del segle XX i que forma part del reconeixement que li va valer el prestigiós Premi Abel al matemàtic hongarès Endre Szemerédi el 2012.

L'anomenat *lema de regularitat* de Szemerédi estableix que, de qualsevol graf prou gran, se'n pot extreure una part estructurada i una part aleatòria. La manera explícita i quantitativa de descriure ambdues parts és el que dona força al teorema.

Sigui $G = (V, E)$ un graf. Per a cada parell de subconjunts disjunts $X, Y \subset V$ denotem per

$$d_G(X, Y) = \frac{e_G(X, Y)}{|X| \cdot |Y|}$$

la densitat d'arestes del parell. Recordem que $e_G(X, Y)$ denota el nombre d'arestes de G amb un vèrtex a X i un a Y, de manera que la densitat $d_G(X, Y)$ és una mesura normalitzada del nombre d'arestes entre X i Y respecte del seu valor màxim $|X| \cdot |Y|$. Si ens concentrem en el subgraf $H = G[X \cup Y]$ induït per $X \cup Y$, aleshores, per a cada $U \subset X$ i $W \subset Y$, es pot mesurar la diferència entre $d_H(U, W)$ i $d_H(X, Y)$. Si les arestes de H estan ben repartides, aleshores aquesta diferència és petita, almenys si U i W són subconjunts prou grans (per a subconjunts petits no podem esperar que la diferència de densitats relatives sigui petita). Per a $\epsilon > 0$ fixat, es diu que el parell (X, Y) és *ϵ-regular* si, per a cada parell de subconjunts $U \subset X$ i $W \subset Y$ tals que $|U| \geq \epsilon|X|$ i $|W| \geq \epsilon|Y|$, se satisfà

$$|d_H(U, W) - d_H(X, Y)| < \epsilon. \tag{1}$$

El que expressa la desigualtat és que un parell ϵ-regular (X, Y) té les arestes repartides de manera quasi aleatòria, en el sentit que la densitat de parells de subconjunts prou grans és pròxima al valor esperat en un graf aleatori amb bipartició $\{X, Y\}$ i densitat d'arestes $d_H(X, Y)$. Amb aquesta terminologia es pot enunciar el lema de regularitat de Szemerédi [15]. Es diu que una partició $V = V_1 \cup \cdots \cup V_r$ d'un conjunt V és una *equipartició* si totes les parts tenen aproximadament el mateix cardinal; més precisament, si $||V_i| - |V_j|| \le 1$ per a cada parell i, j.

TEOREMA 13 (LEMA DE REGULARITAT). *Per a cada $\epsilon > 0$ existeix un $k = k(\epsilon)$ tal que qualsevol graf $G = (V, E)$ admet una equipartició $V = V_1 \cup \cdots \cup V_r$ amb $1/\epsilon < r < k(\epsilon)$ tal que tots els parells (V_i, V_j) són ϵ-regulars llevat com a molt de ϵr^2 parells.*

El lema de regularitat estableix que qualsevol graf prou gran admet una equipartició del conjunt de vèrtexs en un nombre, que depèn només de ϵ i no del graf, tal que gairebé tots els subgrafs induïts pels parells de parts són ϵ-regulars, és a dir, propers a un graf aleatori. Una partició amb aquesta propietat es diu que és una partició *ϵ-regular*. Per a grafs petits o grafs dispersos, es pot prendre la partició de G en vèrtexs aïllats i la descripció estructural resulta trivial (degut a la normalització en la definició de *densitat d'un parell*).

Una de les conseqüències del lema de regularitat en termes de la distància de tall és la següent. Sigui $\{X, Y\}$ un parell ϵ-regular en un graf G i denotem per $G_{X\cup Y}$ el subgraf bipartit de G que té per conjunt de vèrtexs $X \cup Y$ i les arestes de G que uneixen X i Y. Anomenem $\tilde{G}_{X\cup Y}$ el graf bipartit aleatori amb bipartició $\{X, Y\}$ i densitat d'arestes $d_G(X, Y)$. El fet que un parell sigui ϵ-regular es pot expressar en termes de la distància de tall com

$$d_\square(G_{X\cup Y}, \tilde{G}_{X\cup Y}) < \epsilon.$$

Donada una partició $P = (V_1, \ldots, V_r)$ del conjunt de vèrtexs V del graf G, denotem per G_P el graf r-partit amb la mateixa partició tal que cada subgraf $G_P[V_i, V_j]$ és un graf aleatori bipartit amb densitat d'arestes $d_G(V_i, V_j)$ (cada aresta escollida independentment amb probabilitat $p = d_G(V_i, V_j)$). El lema de regularitat implica que, per a cada $\epsilon > 0$, hi ha una equipartició P de V en un nombre de parts que depèn només de ϵ tal que

$$d_\square(G, G_P) < 2\epsilon. \tag{2}$$

En aquesta desigualtat es té en compte que hi pot haver ϵr^2 parells que no són ϵ-regulars, fet que queda absorbit escrivint 2ϵ en lloc de ϵ a la fita superior.

El lema de regularitat ha resultat ser una eina poderosa per a l'anàlisi de grans grafs i té múltiples aplicacions, en particular, en la prova del teorema de Szemerédi sobre l'existència de progressions aritmètiques en conjunts densos d'enters (vegeu, per exemple, [8] o l'article de Lugosi i Serra [13] en aquest mateix *Butlletí*).

Un dels inconvenients del lema de regularitat en les aplicacions, particularment a la teoria de límits de grafs, és la relació entre ϵ i k, ja que la fita superior per a k no pot ser millor que una torre exponencial $2^{2^{\cdot^{\cdot^{2}}}}$ d'alçada $1/\epsilon$, tal com Gowers ([4]) va provar.

Alan Frieze i Ravi Kannan ([3]) van donar una versió més dèbil del lema de regularitat que recull la desigualtat (2) i que dona una relació molt més eficient entre k i ϵ. Aquesta versió és adequada per a la seva aplicació als límits de grafs. En termes de la distància de tall, la versió dèbil del lema de regularitat de Frieze i Kannan s'enuncia de la manera següent.

Teorema 14 (Lema dèbil de regularitat). *Per a cada enter positiu k, qualsevol graf G admet una equipartició P en k parts tal que*

$$d_\square(G, G_P) \leq \frac{4}{\sqrt{k}}.$$

Frieze i Kannan van estendre el lema dèbil de regularitat també a l'espai dels grafons. Per a aquesta versió s'aproxima un grafó per un grafó esgraonat. Donada una partició $P = \{S_1, \ldots, S_k\}$ de $[0,1]$ en k conjunts mesurables, un grafó esgraonat és un grafó que és constant sobre $S_i \times S_j$ per a cada i, j. Per exemple, donat un graf G, el grafó W_G que s'ha definit a la secció 2 és un grafó esgraonat (que només pren valors a $\{0,1\}$). Per a un grafó W i una partició P de $[0,1]$ en k conjunts mesurables de mesura no nul·la, denotem per W_P el grafó esgraonat que sobre cada conjunt $S_i \times S_j$ pren el valor mitjà de W sobre aquest conjunt:

$$W_P|_{S_i \times S_j} \equiv \frac{1}{\lambda(S_i)\lambda(S_j)} \int_{S_i \times S_j} W(x,y)\, dx\, dy,$$

on λ denota la mesura de Lebesgue.

Teorema 15 (Lema dèbil de regularitat). *Per a cada enter positiu k i qualsevol grafó W hi ha una partició P de $[0,1]$ en k conjunts mesurables de mesura no nul·la tal que*

$$d_\square(W, W_P) \leq \frac{4}{\sqrt{k}}.$$

Una partició P de $[0,1]$ tal que $d_\square(W, W_P) < \epsilon$ s'anomena una *partició ϵ-regular de W*. L'equivalència de nocions amb el cas dels grafs és clara, i l'enunciat en el cas dels grafons resulta més transparent.

9 Lema de regularitat i densitat d'immersions

La connexió entre les successions $t(\cdot, G)$ i $t(\cdot, W)$, basades a comptar subgrafs en grafs o grafons, i el lema de regularitat es fonamenten en l'anomenat *lema de comptatge*, una de les aplicacions més notables del lema de regularitat. El lema de comptatge dona una aproximació del nombre de còpies de certs subgrafs de H en un graf G a través de particions regulars. Si $P = \{V_1, \ldots, V_r\}$ és una partició ϵ-regular de G, denotem per $\tilde{G}_P$ el graf aleatori r-partit que, per a cada parell de vèrtexs x, y en parts diferents V_i, V_j, conté l'aresta $\{x,y\}$

amb probabilitat $d(V_i, V_j)$. El nombre esperat de còpies d'un subgraf H de G/P a $\tilde{G}_P$ és aleshores

$$\prod_{ij\in E(H)} d_G(V_i, V_j) \prod_{i=1}^{r} |V_r|.$$

La regularitat de la partició permet veure que el nombre de còpies de H en G és aproximadament aquest nombre esperat. Més precisament, el nombre $\mathrm{ind}(H, G)$ de còpies d'un subgraf H de G/P a G satisfà

$$\left|\mathrm{ind}(H, G) - \prod_{ij\in E(H)} d_G(V_i, V_j) \prod_{i=1}^{r} |V_r|\right| \leq \epsilon e(H) \prod_{i=1}^{r} |V_r|,$$

on $e(H)$ és el nombre d'arestes de H. En termes de la densitat d'immersions, aquest lema es pot reformular de la manera següent. D'acord amb el lema dèbil de regularitat (teorema 14), per a cada $\epsilon > 0$ hi ha una partició P de G tal que $d_\square(G, G_P) < \epsilon$ amb un nombre de parts $k = 1/\epsilon^2$. El següent enunciat dona, doncs, una forma en cert sentit més general del lema de comptatge.

Teorema 16 (Lema de comptatge). *Sigui H un graf. Per a dos grafs G, G' se satisfà*

$$|t(H, G) - t(H, G')| \leq e(H) d_\square(G, G').$$

El lema de comptatge anterior s'estén de manera natural a densitat d'immersions en grafons (Lovász i Szegedy [12]).

Teorema 17 (Lema de comptatge). *Sigui H un graf. Per a dos grafons W, W' se satisfà*

$$|t(H, W) - t(H, W')| \leq e(H) d_\square(W, W').$$

Aquest darrer resultat es pot expressar dient que, per a un graf H fixat, l'aplicació

$$\begin{aligned} \phi\colon \mathcal{W}_0 &\to [0,1] \\ W &\mapsto t(H, W) \end{aligned}$$

és de Lispschitz en l'espai mètric $(\mathcal{W}_0, d_\square)$. A més, es pot provar que l'aplicació és injectiva i té inversa contínua.

Finalment, es pot enunciar el teorema, que és el nostre objectiu final.

Teorema 18 (Lovász i Szegedy [12]). *Sigui G_n una successió convergent de grafs. Existeix un grafó W tal que la successió convergeix a W.*

Prova. Pel teorema 7, l'espai dels grafons $(\mathcal{W}_0, d_\square)$ amb la distància de tall és compacte. D'altra banda, si $\mathcal{G}$ denota l'espai dels grafs finits, l'espai de seqüències $\{t(\cdot, G) : G \in \mathcal{G}\}$ és isomorf a $[0,1]^{\mathcal{G}\times\mathcal{G}}$, que també és compacte, pel teorema de Tijonov. L'aplicació $W \mapsto (\cdot, W)$ és contínua pel teorema 17, i té inversa contínua. Si G_n és una successió convergent de grafs, aleshores la successió $t(H, G_n)$ és convergent per a cada H i $t(H, W_{G_n})$ també ho és. Per la continuïtat de l'aplicació anterior, la successió W_{G_n} també és convergent a $(\mathcal{W}_0, d_\square)$. $\square$

Recapitulant el que s'ha exposat fins aquí, hem donat una definició de *convergència de successions de grafs* en termes de nombres d'immersions i hem definit una distància en l'espai dels grafs, la distància de tall, segons la qual una successió és convergent si, i només si, és de Cauchy respecte de la distància de tall. D'altra banda, hem definit la noció de convergència a un grafó i hem vist que cada grafó és el límit d'una successió de grafs i que la convergència es pot descriure en termes de la distància de tall. Finalment, hem constatat que cada grafó és el límit d'una successió de grafs. D'aquesta manera hi ha un marc sòlid per a estudiar límits de grafs densos.

10 Per acabar

En aquest article s'ha introduït la noció de límit d'una successió de grafs i els grafons com a objectes límit. Aquest és només l'inici d'una història molt més llarga que posa en acció aquesta noció per a obtenir resultats sobre l'anàlisi de grans xarxes. D'aquesta història se n'han escrit ja uns quants capítols, però en queden molts més per escriure. Els grafons són objectes en els quals es poden aplicar les poderoses eines de l'anàlisi matemàtica. Com passa sovint, molts problemes són més tractables en el continu que en el discret, i els resultats són sovint més compactes i transparents.

D'altra banda, els grafons obren un nombre d'interrogants que segurament no es plantejarien des de la perspectiva discreta, problemes de naturalesa topològica i analítica, alguns dels quals han trobat fins ara respostes inesperades, que han enriquit l'univers dels objectes límit.

Com ja hem comentat, la noció de grafó només resulta útil quan es consideren successions de grafs densos. En el cas dels grafs dispersos, amb un nombre subquadràtic d'arestes, la noció de límit s'ha de reformular d'una altra manera, en particular, per l'absència del lema de regularitat de Szemerédi per a aquesta classe de grafs. En aquest sentit, s'han proposat diverses alternatives per a la noció de convergència, en concret per a la classe de grafs en què el grau màxim dels vèrtexs és fitat. Aquesta classe no inclou, però, exemples importants de grafs, com els grafs plans o molts dels models de grans xarxes, en els quals la successió de graus dels vèrtexs segueix una llei potencial. En aquest sentit, la teoria està en plena expansió i s'aborda des d'angles molt diferents, des de la perspectiva de la probabilitat, de la lògica, de la topologia, de l'anàlisi o de l'àlgebra, cadascuna aportant eines i punts de vista substancials.

Els desenvolupaments descrits en els paràgrafs anteriors estan exposats al text de Lovász [11] i alguns dels avenços més recents estan reflectits a l'article de Grzesik i Král' [5], on el lector interessat trobarà material i referències abundants per a continuar aprofundint en aquesta apassionant aventura matemàtica.

La importància científica de l'anàlisi matemàtica de grans xarxes ha estat reconeguda per l'*European Research Council*, que, en la seva convocatòria del 2018 en el programa «ERC Synergy Grants», finança el projecte *Dynamics and Structure of Networks*, liderat per Lászó Lovász, László Barabássy i Jaroslav

Nešetřil, l'únic projecte purament matemàtic en aquest programa estrella de la Unió Europea. De ben segur que en els propers anys la teoria dels límits de grafs tindrà una evolució substancial i, qui sap, es podrà fer realitat el somni de Lovász de contribuïr a entendre millor no solament les grans xarxes humanes, biològiques o tecnològiques, sino també el món que ens envolta i la seva immensa xarxa d'interaccions.

Referències

[1] Biggs, N. L.; Lloyd, E. K.; Wilson, R. J. *Graph Theory. 1736–1936*. 2a ed. Nova York: The Clarendon Press, Oxford University Press, 1986.

[2] Diestel, R. *Graph Theory*. 5a ed. Berlín: Springer, 2018. (Graduate Texts in Mathematics; 173)

[3] Frieze, A.; Kannan, R. «Quick approximation to matrices and applications». *Combinatorica*, 19 (2) (1999), 175–220.

[4] Gowers, W. T. «Lower bounds of tower type for Szemerédi's uniformity lemma». *Geom. Funct. Anal.*, 7 (2) (1997), 322–337.

[5] Grzesik, A.; Král', D. «Analytic representations of large graphs». A: *Surveys in Combinatorics 2019*. Cambridge: Cambridge University Press, 2019, 57–88. (London Math. Soc. Lecture Note Ser.; 456)

[6] Hell, P.; Nešetřil, J. «On the complexity of H-coloring». *J. Combin. Theory Ser. B*, 48 (1) (1990), 92–110.

[7] Hell, P.; Nešetřil, J. *Graphs and Homomorphisms*. Oxford: Oxford University Press, 2004. (Oxford Lecture Series in Mathematics and its Applications; 28)

[8] Komlós, J.; Shokoufandeh, A.; Simonovits, M.; Szemerédi, E. «The regularity lemma and its applications in graph theory». A: *Theoretical Aspects of Computer Science*. Berlín: Springer, 2002, 84–112. (Lecture Notes in Comput. Sci.; 2292)

[9] Král', D. «Graph limits and extremal graph theory». Curs a São Paulo School of Advanced Science on Algorithms, Combinatorics and Optimization (2016).

[10] Lovász, L. «Operations with structures». *Acta Math. Acad. Sci. Hungar.*, 18 (1967), 321–328.

[11] Lovász, L. *Large Networks and Graph Limits*. Providence, RI: American Mathematical Society, 2012. (American Mathematical Society Colloquium Publications; 60)

[12] Lovász, L.; Szegedy, B. «Limits of dense graph sequences». *J. Combin. Theory Ser. B*, 96 (6) (2006), 933–957.

[13] Lugosi, G.; Serra, O. «Endre Szemerédi, Premi Abel 2012». *Butlletí de la Societat Catalana de Matemàtiques*, 28 (1) (2013), 87–115.

[14] McDiarmid, C. «Concentration». A: *Probabilistic Methods for Algorithmic Discrete Mathematics*. Berlín: Springer, 1998, 195–248. (Algorithms Combin.; 16)

[15] Szemerédi, E. «Regular partitions of graphs». A: *Problèmes combinatoires et théorie des graphes* (Colloq. Internat. CNRS, Univ. Orsay, Orsay, 1976). París: CNRS, 1978, 399–401. (Colloq. Internat. CNRS; 260)

Departament de matemàtiques
Universitat Politècnica de Catalunya
08034 Barcelona
oriol.serra@upc.edu