Axioma d'elecció i lema de Zorn

Jordi Quer

26 de març de 2020

Índex

1 Repàs de famílies, conjunts i ordres

Quan es vol treballar amb una col·lecció formada per alguns elements d'un conjunt X es poden considerar els elements que pertanyen a un subconjunt $S \subseteq X$. En algunes situacions és convenient considerar col·leccions d'elements de X en que poden aparèixer elements repetits i en que cada element porta una "etiqueta" que l'identifica. Per aconseguir això, l'objecte matemàtic adequat s'anomena *família* o *família indexada* d'elements del conjunt X.

Famílies d'elements d'un conjunt. Donat un conjunt X, una *família* d'elements de X amb conjunt d'índexs I és una col·lecció $(x_i)_{i \in I}$ d'elements x_i del conjunt X, un per a cada índex $i \in I$. Noti's que els elements de X poden aparèixer repetits en una família: pot ser que $x_i = x_j$ per a índexs diferents $i \neq j$. També s'anomenen I*-tuples*.

Típicament, les famílies finites s'indexen amb el conjunt $I = \{1, 2, \ldots, n\}$ o un de semblant, de manera que la família és la n-tupla o vector $(x_1, x_2, \ldots, x_n)$. En particular, l'ordre natural del conjunt I dóna una ordenació dels n elements de la família. Les famílies numerables s'acostumen a indexar amb el conjunt $\mathbb{N} = \{1, 2, 3, \ldots\}$: són les *successions* $(x_n)_{n \geqslant 1}$ d'elements de X.

Donar una família d'elements de X indexada pel conjunt I és el mateix que donar una aplicació $I \to X$ que assigna a cada índex $i \in I$ un element $x_i \in X$. Per tant, el concepte de família d'elements d'un conjunt no és nou, i es tracta només d'una notació que és molt pràctica en determinades situacions, per treballar amb un objecte matemàtic ben conegut: una aplicació entre dos conjunts.

Una *subfamília* d'una família $(x_i)_{i \in I}$ d'elements de X és una altra família $(y_j)_{j \in J}$ d'elements de X tal que el seu conjunt d'índexs és un subconjunt del de la primera, o sigui $J \subseteq I$, i, per a cada índex d'aquest subconjunt J, l'element de totes dues famílies corresponent és el mateix: $y_j = x_j$ per a tot $j \in J$. En termes de l'aplicació $I \to X$ una subfamília és simplement la restricció de l'aplicació a un subconjunt $J \subseteq I$.

Tota família $(x_i)_{i \in I}$ determina *un únic* subconjunt de X: el subconjunt

$$S = \{x_i : i \in I\} = \{x \in X : \exists i \in I, x_i = x\}$$

format pels elements de X que són termes de la família. És simplement el conjunt imatge de l'aplicació $i \mapsto x_i : I \to X$.

En canvi el mateix conjunt S pot provenir de moltes famílies diferents: totes les que l'aplicació $I \to X$ corresponent té per imatge el conjunt S. En particular, agafant un conjunt d'índexs I amb el mateix cardinal que S i una bijecció $I \to S \subseteq X$ es té una família tal que els elements del conjunt apareixen tots com a un terme i només un. D'aquesta manera, si cal, un conjunt es pot convertir en una família sense que hi hagi duplicitats en els elements.

Definició 1 (Ordre) *Un* ordre *a un conjunt X és una relació, que acostuma a denotar-se $x \leqslant y$, que sigui reflexiva, antisimètrica i transitiva. Un* conjunt ordenat *és un conjunt on hi ha definit un ordre; pot indicar-se com un parell ordenat $(X, \leqslant)$.*

La restricció d'un ordre a un subconjunt és també un ordre i per tant tot subconjunt d'un conjunt ordenat és també un conjunt ordenat amb l'ordre heretat del conjunt.

Dos elements x, y d'un conjunt ordenat es diuen *comparables* si $x \leqslant y$ o bé $y \leqslant x$. Per exemple, en el conjunt dels nombres reals amb la relació d'ordre habitual, dos elements qualsevol sempre són comparables: un dels dos és sempre menor o igual que l'altre. En canvi, al conjunt $A = \mathscr{P}(A)$ de les parts d'un conjunt A, ordenat segons la relació d'inclusió de conjunts, dos subconjunts diferents $\{a\}$ i $\{b\}$ que contenen cadascun un únic element no són comparables.

Definició 2 (Ordre total) *Un ordre es diu* total *si tot parell d'elements del conjunt són comparables. Moltes vegades a un ordre qualsevol se li diu* ordre parcial *per fer èmfasi en que no necessàriament compleix la condició de ser un ordre total. Un conjunt amb un ordre total es diu* totalment ordenat*.*

És clar que tot subconjunt d'un conjunt totalment ordenat és totalment ordenat i que tot conjunt totalment ordenat finit té element mínim i element màxim (també es poden dir primer i últim).

Definició 3 (Cadena) *Es diu que un subconjunt $C \subseteq X$ és una* cadena *de X si l'ordre a C heretat de l'ordre de X és un ordre total en el subconjunt C. O sigui, una cadena d'un conjunt ordenat és un subconjunt totalment ordenat.*

Per exemple, el subconjunt $\{\{0, 1, 2, \ldots, n\} : n \geqslant 1\}$ és una cadena dins de $\mathscr{P}(\mathbb{Z})$ però el subconjunt $\{\{0, n\} : n \geqslant 1\}$ no ho és. Naturalment, tot subconjunt d'un conjunt totalment ordenat és una cadena. És clar que tot subconjunt finit d'un conjunt totalment ordenat té un element màxim.

Definició 4 (Elements especials) *Donat un subconjunt $S \subseteq X$ d'un conjunt ordenat,*

- $x \in X$ *és una* fita superior *de S si és més gran o igual que tots els elements de S;*
- $x \in X$ *és un* suprem *de S si és fita superior i és menor o igual que totes les fites superiors;*
- $s \in S$ *és un* màxim *de S si és més gran o igual que tots els elements de S;*
- $s \in S$ *és un element* maximal *de S si no hi ha cap element a S estrictament més gran.*

De manera anàloga es defineixen fita inferior*,* ínfim*,* mínim *i element* minimal*.*

El màxim i el suprem d'un subconjunt S, si existeixen, són únics; en canvi un subconjunt pot tenir moltes fites superiors i molts elements maximals diferents. De vegades al mínim d'un conjunt se li diu *primer element*.

Definició 5 (Bon ordre) *Un conjunt ordenat X es diu* ben ordenat *si tot subconjunt no buit té mínim. En aquest cas es diu que l'ordre del conjunt X és un* bon ordre*.*

Tot conjunt ben ordenat està totalment ordenat ja que donats dos elements del conjunt el subconjunt no buit format per aquests dos elements té mínim, i per tant un dels dos ha de ser menor o igual que l'altre.

Definició 6 (Principi de bona ordenació) *Es coneix amb el nom de* principi de bona ordenació *l'afirmació que diu que en tot conjunt es pot definir un ordre que és un bon ordre.*

La teoria axiomàtica de conjunts estàndard, amb el sistema d'axiomes de Zermelo-Fraenkel, inclou l'*axioma d'elecció*. Aquest axioma assegura que el producte cartesià d'una família de conjunts no buits conté algun element: donada una família de conjunts $(X_i)_{i \in I}$ amb $X_i \neq \varnothing$ per a tot índex i aleshores existeix una família $(x_i)_{i \in I}$ amb $x_i \in X_i$ per a tot índex i. És a dir, la família $(x_i)_{i \in I}$ és un element del producte cartesià de conjunts $\prod_{i \in I} X_i$.

Dit d'una altra manera, donats conjunts X_i qualsevol, sempre "hi ha una manera de triar un element x_i en cadascun dels conjunts X_i".

Definició 7 (Axioma d'elecció) *L'*axioma d'elecció *és un axioma de la teoria de conjunts que diu el següent: per a tota família de conjunts no buits $(X_i)_{i \in I}$ existeix una aplicació $f \colon I \to \bigcup_{i \in I} X_i$ tal que $f(i) \in X_i$ per a tot $i \in I$.*

2 Lema de Zorn i axiomes equivalents

Definició 8 (Lema de Zorn) *Es coneix com a* lema de Zorn *l'afirmació següent: Tot conjunt ordenat no buit tal que tota cadena té alguna fita superior conté elements maximals.*

O sigui, donat un conjunt X amb un ordre parcial $\leqslant$ tal que per a tot subconjunt $C \subseteq X$ que sigui una cadena (estigui totalment ordenat) existeix un element $c \in X$ tal que $x \leqslant c\, \forall\, x \in C$, aleshores existeix un element $m \in X$ tal que $m \leqslant x \Rightarrow m = x$, $\forall\, x \in C$.

De vegades un conjunt que satisfaci la hipòtesi del lema de Zorn s'anomena *conjunt inductiu.* Per exemple, els subconjunts finits d'un conjunt infinit, ordenats per inclusió, formen un conjunt no inductiu. En canvi, els subconjunts linealment independents d'un espai vectorial, ordenats per inclusió, formen un conjunt inductiu (i aplicant el lema de Zorn es dedueix immediatament que tot espai vectorial té bases: elements maximals del conjunt dels subconjunts de vectors linealment independents).

Equivalència. Els tres enunciats de les definicions 6, 7 i 8, a pesar que per tradició s'anomenen de manera diferent (principi, axioma i lema), són tots tres enunciats equivalents. Són de fet un axioma, que forma part de totes les teories matemàtiques, només amb algunes excepcions lligades sobretot a estudis de lògica i fonamentació.

Per aplicar-lo en la pràctica la versió més útil és el lema de Zorn. Molts resultats estàndard que es fan servir contínuament en matemàtiques requereixen usar aquest axioma (de fet, molts d'ells són enunciats equivalent). En àlgebra lineal l'existència de bases i el fet que dues bases tinguin el mateix cardinal (teorema de Steinitz); en topologia el teorema de Tychonoff, que assegura que el producte d'espais compactes és compacte, i que és el motiu pel qual ara s'està estudiant aquest tema; en àlgebra el fet que tot anell no trivial té ideals maximals; en teoria de Galois l'existència de clausures algebraiques i l'aixecament d'immersions; etc.

A continuació es dóna una demostració d'aquesta equivalència en el teorema 11, després d'una definició i un lema previ.

Definició 9 (Segment inicial) *Si S és un subconjunt d'un conjunt ordenat X, per a cada element $s \in S$ el subconjunt $S_{<s} = \{x \in S : x < s\}$ s'anomena* segment inicial *de S.*

Lema 10 *Sigui $(S_i)_{i\in I}$ una família de subconjunts ben ordenats d'un conjunt ordenat. Suposi's que donats dos elements diferents $S_i \neq S_j$, un d'ells és sempre un segment inicial de l'altre. Aleshores la reunió $S = \bigcup_{i\in I} S_i$ és un conjunt ben ordenat.*

Prova: Sigui $T \subseteq S$ un subconjunt no buit. sigui S_j un dels subconjunts de la família que tingui intersecció no buida amb T. Sigui $m = \min(S_j \cap T)$, que existeix gràcies a la bona ordenació de S_j. Es vol veure que $m = \min T$. Tot element $x \in T$ pertany a algun dels subconjunts de la família. Sigui $x \in S_i$. Si $x \in S_j$ aleshores ja es té $m \leqslant y$. Si $x \notin S_j$ aleshores no pot ser $S_i \subseteq S_j$ i per tant ha de ser $S_j \subseteq S_i$; a més, el primer d'aquests subconjunts ha de ser un segment inferior del segon, i per tant $m < x$. □

Teorema 11 *L'axioma d'elecció, el principi de bona ordenació i el lema de Zorn són tres enunciats equivalents.*

Prova: *Bona ordenació $\Rightarrow$ Elecció.* Donada una família de conjunts no buits $(X_i)_{i\in I}$ es considera la seva reunió $X = \bigcup_{i\in I} X_i$. Per hipòtesi existeix un bon ordre $\leqslant$ al conjunt

X. Per a cada $i \in I$ es defineix x_i com el mínim element del subconjunt no buit $X_i \subseteq X$. Aleshores $(x_i)_{i\in I}$ és un element del producte cartesià $\prod_{i\in I} X_i$.

Elecció $\Rightarrow$ *Zorn.* Sigui X un conjunt ordenat no buit tal que tota cadena té alguna fita superior. Per demostrar l'implicació se suposa que X no té cap element maximal i s'ha d'arribar a una contradicció.

Que X no tingui elements maximals vol dir que tot element de X té algun element estrictament més gran. Aleshores tota cadena $C \subseteq X$ té fites superiors que no pertanyen a la cadena, ja que donada una fita superior qualsevol, que pot pertànyer o no a la cadena, tot element de X estrictament més gran que aquesta fita superior també és fita superior i no pertany a la cadena. Sigui $\mathscr{C}$ el conjunt de les cadenes de X. Per a cada $C \in \mathscr{C}$ es defineix el subconjunt $X_C \subseteq X$ com el conjunt de les fites superiors de C que no pertanyen a C, que és un conjunt no buit. L'axioma d'elecció garanteix l'existència d'una aplicació $f\colon \mathscr{C} \to \bigcup_{C\in\mathscr{C}} X_C \subseteq X$ tal que per a cada cadena $C \in \mathscr{C}$ la imatge $f(C)$ és una fita superior de C que no pertany a C.

Es dirà que un subconjunt A de X és *admissible* (respecte l'aplicació f) si està ben ordenat (per tant, és una cadena) i per a tot segment inicial $A_{<a} = \{x \in A : x < a\}$ corresponent a un $a \in A$ es té $f(A_{<a}) = a$.

Primerament es veurà que donats dos subconjunts admissibles diferents $A \neq B$ aleshores un d'ells és un segment inicial de l'altre. Suposi's que $B \not\subseteq A$. S'ha de demostrar que A és segment inicial de B. Com que $B \setminus A \subseteq B$ és no buit, per ser B ben ordenat existeix $b = \min(B \setminus A)$. Per tant, $B_{<b} = \{x \in B : x < b\} \subseteq A$. Es veurà que aquesta inclusió és una igualtat per reducció a l'absurd. Suposi's que la inclusió es estricta. Per ser A ben ordenat existeix $a = \min(A \setminus B_{<b})$. Aleshores $A_{<a} \subseteq B_{<b}$. Com que $B \not\subseteq A$ també es té $B \not\subseteq A_{<a}$ i per ser B ben ordenat existeix $c = \min(B \setminus A_{<a})$ i es té una cadena d'inclusions:

$$B_{<c} \subseteq A_{<a} \subseteq B_{<b} \subset A.$$

Els dos primers subconjunts de la cadena són iguals. En efecte, si $c = b$ el primer i el tercer conjunt són iguals i per tant els dos primers també, i si $c < b$ aleshores $c \in B_{<b} \Rightarrow c \in A$, i com que $c \notin A_{<a}$ ha de ser $c \geqslant a$, d'on resulta que $A_{<a} \subseteq A_{<c} \cap B_{<b} = B_{<c}$ i també $B_{<c} = A_{<a}$. Aleshores $c = f(B_{<c}) = f(A_{<a}) = a \in A$. Com que $b \notin A$ ha de ser $c \neq b$ i per tant $c < b$, d'on resulta que $a = c \in B_{<b}$, que és una contradicció amb la definició de a Per tant l'element a no es pot construir i i això assegura la igualtat $A = B_{<b}$. Així, efectivament en cada parell de subconjunts admissibles diferents un dels dos ha de ser un segment inicial de l'altre.

Sigui $\mathscr{A} \subseteq \mathscr{P}(X)$ el subconjunt format per tots els conjunts admissibles, ordenat amb l'ordre definit per la inclusió. Sigui $Y = \bigcup_{A\in\mathscr{A}} A$ la reunió de tots aquests subconjunts. Es vol demostrar que Y és admissible.

El lema 10 assegura que el conjunt Y està ben ordenat ja que és reunió de subconjunts ben ordenats tals que de dos de diferents sempre n'hi ha un que és un segment inicial de l'altre. Per tant només s'ha de comprovar la condició $f(Y_{<y}) = y$ per a tot $y \in Y$. Per a això es veurà primerament que si $y \in A$ per a un subconjunt admissible $A \in \mathscr{A}$ aleshores $Y_{<y} = A_{<y}$. La inclusió $A \subseteq Y$ implica que $A_{<y} \subseteq Y_{<y}$. Per veure la recíproca sigui $x \in Y_{<y}$. Sigui B un conjunt admissible que contingui x. Si $B \subseteq A$ aleshores

$x \in A_{<y}$. Altrament $A = B_{<b}$ és un segment inferior de B. Per tant $y \in A \Rightarrow y < b$ i $x < y \Rightarrow x \in B_{<b} = A \Rightarrow x \in A_{<y}$. A partir de la igualtat $Y_{<y} = A_{<y}$, i aplicant que A admissible, es té $f(Y_{<y}) = f(A_{<y}) = y$. Amb això s'acaba la prova de que Y és admissible.

La contradicció s'obté considerant el conjunt $Y \cup \{f(Y)\}$. Aquest conjunt conté estrictament el conjunt Y i és admissible, i per tant pertany a $\mathscr{A}$ de manera que hauria d'estar contingut a Y, però no ho està ja que $f(Y) \notin Y$ per definició de l'aplicació f.

Zorn $\Rightarrow$ Bona ordenació. Donat un conjunt no buit A, es considera el subconjunt X format pels parells $(S, \leqslant_S)$ d'un subconjunt $S \subseteq A$ i un bon ordre $\leqslant_S$ en aquest conjunt S. Al conjunt X es defineix un ordre parcial posant $(S, \leqslant_S) \leqslant (T, \leqslant_T)$ si el primer és un segment inferior del segon en el sentit següent:

1. $S \subseteq T$,
2. l'ordre $\leqslant_T$ estén l'ordre $\leqslant_S$ (o sigui, $a \leqslant_T b \Leftrightarrow a \leqslant_S b \;\forall\, a, b \in S$), i
3. si $S \neq T$ aleshores $S = T_{<m}$ amb $m = \min(T \smallsetminus S)$.

Es comprova immediatament que aquesta és una relació d'ordre.

Donada una cadena $(S_i, \leqslant_i)$ d'elements de X es considera el conjunt reunió $S = \bigcup S_i$ i, en aquest conjunt, l'ordre $\leqslant_S$ que estén els ordres de cada S_i. És a dir, donats dos elements $a, b \in S$ sigui $a \in S_i$ i $b \in S_j$; com que un d'aquests dos subconjunts està contingut a l'altre, suposi's per exemple que $S_i \subseteq S_j$, aleshores $a, b \in S_j$ i es defineix $a \leqslant_S b \Leftrightarrow a \leqslant_j b$. Aquesta relació està ben definida i és d'ordre gràcies a que els conjunts $(S_i, \leqslant_i)$ formen una cadena i els ordres dels conjunts més grans estenen els ordres dels conjunts més petits. El lema 10 aplicat als subconjunts $S_i \subseteq S$ assegura que el conjunt S està ben ordenat ja que en dos subconjunts diferents l'un és sempre segment inicial de l'altre. Per tant, $(S, \leqslant_S) \in X$. Es vol veure que $(S_i, \leqslant_i) \leqslant (S, \leqslant_S)$. La inclusió $S_i \subseteq S$ i el fet que $\leqslant_S$ estén $\leqslant_i$ són immediats per la definició de S i $\leqslant_S$. Per tant només falta veure que cada $S_i \neq S$ és un segment inicial de S; concretament, si $m = \min(S \smallsetminus S_i)$ s'ha de veure que $S_i = S_{<m}$. Sigui S_m un subconjunt que contingui m. Els subconjunts S_i i S_m són diferents i per tant un és segment inicial de l'altre; clarament ha de ser $S_i = (S_m)_{<s}$i com que $m \notin S_i$ cal que $s \leqslant m$. Es dedueix que $S_i = (S_m)_{<s} \subseteq (S_m)_{<m} \subseteq S_{<m}$. Donat $x \in S_{<m} \subseteq S$ ha de ser $x \in S_i$ ja que si no $m \leqslant x$.

Aplicant el lema de Zorn, el conjunt X té elements maximals. Sigui $(M, \leqslant_M)$ un element maximal. Es vol demostrar que $M = A$ per reducció a l'absurd. Suposi's que hi hagués un element $a \in A \smallsetminus M$. Aleshores el conjunt $N = M \cup \{a\}$ admet el bon ordre $\leqslant_N$ definit posant $x \leqslant_N y \Leftrightarrow x \leqslant_N y$ per a tot parell d'elements $x, y \in M$ i afegint les desigualtats $x <_N a$ per a tot $x \in M$. A més M és clarament un segment inferior de N. Per tant seria $(M, \leqslant_M) < (N, \leqslant_N)$ i això contradiu la maximalitat del primer.

Per tant, el conjunt A admet una bona ordenació. $\square$