Submodular Functions, Optimization, and Applications to Machine Learning

— Spring Quarter, Lecture 14 —

http://j.ee.washington.edu/~bilmes/classes/ee596b_spring_2014/

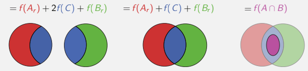
Prof. Jeff Bilmes

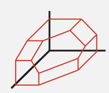
University of Washington, Seattle
Department of Electrical Engineering
http://melodi.ee.washington.edu/~bilmes

May 18th, 2014

$f(A) + f(B) \geq f(A \cup B) + f(A \cap B)$

$= f(A_r) + 2f(C) + f(B_r) \quad = f(A_r) + f(C) + f(B_r) \quad = f(A \cap B)$





Cumulative Outstanding Reading

- Good references for today: Birkhoff, "Lattice Theory", 1967.

Announcements, Assignments, and Reminders

- This is a special extra lecture on lattices that is not given during regular lecture time.

Class Road Map - IT-I

- L1 (3/31): Motivation, Applications, & Basic Definitions
- L2: (4/2): Applications, Basic Definitions, Properties
- L3: More examples and properties (e.g., closure properties), and examples, spanning trees
- L4: proofs of equivalent definitions, independence, start matroids
- L5: matroids, basic definitions and examples
- L6: More on matroids, System of Distinct Reps, Transversals, Transversal Matroid, Matroid and representation
- L7: Dual Matroids, other matroid properties, Combinatorial Geometries
- L8: Combinatorial Geometries, matroids and greedy, Polyhedra, Matroid Polytopes,
- L9: From Matroid Polytopes to Polymatroids.
- L10: Polymatroids and Submodularity
- L11: More properties of polymatroids, SFM special cases
- L12: polymatroid properties, extreme points polymatroids,
- L13: sat, dep, supp, exchange capacity, examples
- L14: Lattice theory: partially ordered sets; lattices; distributive, modular, submodular, and boolean lattices; ideals and join irreducibles.
- L15:
- L16:
- L17:
- L18:
- L19:
- L20:

Finals Week: June 9th-13th, 2014.

Outline

- We're next going to study lattices and submodular functions.
- In doing so, we'll better be able to understand certain properties of polymatroidal extreme points and ultimately SFM.

Partially ordered set

- A partially ordered set (poset) is a set of objects with an order.
- Set of objects V and a binary relation $\preceq$ which can be read as "is contained in" or "is part of" or "is less than or equal to".
- For any $x, y \in V$, we may ask is $x \preceq y$ which is either true or false.
- In a poset, for any $x, y, z \in V$ the following conditions hold (by definition):

For all $x, x \preceq x$.	(Reflexive)	(P1.)
If $x \preceq y$ and $y \preceq x$, then $x = y$	(Antisymmetriy)	(P2.)
If $x \preceq y$ and $y \preceq z$, then $x \preceq z$.	(Transitivity)	(P3.)

- We can use the above to get other operators as well such as "less than" via $x \preceq y$ and $x \neq y$ implies $x \prec y$. Also, we get $x \succ y$ if not $x \preceq y$. And $x \succeq y$ is read "x contains y". And so on.

Partially ordered set

- A partially ordered set (poset) is a set of objects with an order.
- In a poset, for any $x, y, z \in V$ the following conditions hold (by definition):

For all $x, x \preceq x$.	(Reflexive)	(P1.)
If $x \preceq y$ and $y \preceq x$, then $x = y$	(Antisymmetriy)	(P2.)
If $x \preceq y$ and $y \preceq z$, then $x \preceq z$.	(Transitivity)	(P3.)

- The order $n(P)$ of a poset P is meant the (cardinal) number of its elements.

Partially ordered set

- Given two elements, we need not have either $x \preceq y$ or $y \preceq x$ be true, i.e., these elements might not be comparable. If for all $x, y \in V$ we have $x \preceq y$ or $y \preceq x$ then the poset is **totally ordered**.
- There may exist only one element x which satisfies $x \preceq y$ for all y. Since if $x \preceq y$ for all y and $z \preceq y$ for all y then $z \preceq x$ and $x \preceq z$ implying $x = z$. If it exists, we can name this element 0 (zero). The dual maximal element is called 1 (one).
- We define a set of elements $x_1, x_2, \ldots, x_n$ as a chain if $x_1 \preceq x_2 \preceq \cdots \preceq x_n$, which means $x_1 \preceq x_2$ and $x_2 \preceq x_3$ and $\ldots x_{n-1} \preceq x_n$. While we normally thing of the elements of a chain as distinct they need not be. The length of a chain of n elements is $n - 1$.

Partially ordered set

Example 14.3.1

Let $V = \mathbb{Z}^+$ be the set of positive integers and let $x \preceq y$ mean that x is less than y in the usual sense. Then we have a poset that is actually totally ordered.

Example 14.3.2

Let V consist of all real single-valued functions $f(x)$ defined on the closed interval $[-1, 1]$, and let $g \leq f$ mean that $g(x) \leq f(x)$ for all $x \in [-1, 1]$. Again poset, but not total order.

- Any subset of a poset is a poset. If $S \subseteq V$ than for $x, y \in S$, $x \preceq y$ is the same as taken from V, but we just restrict the items to S.
- Any subset of a chain is a chain.
- Two posets V_1 and V_2 are isomorphic if there is an isomorphism between them (i.e., a 1-1 order preserving (isotone) function that has an order preserving inverse). We write that two posets U and V are isomorphic by $U \simeq V$.

Partially ordered set

- **duality**. The dual poset is formed by exchanging $\preceq$ with $\succeq$. This is called the converse of a partial ordering. The converse of a PO is also a PO. We write the dual of V as V^d. U and V are dually isomorphic if $U = V^d$ or equivalently $V = U^d$. When $U = U^d$ then U is self-dual.

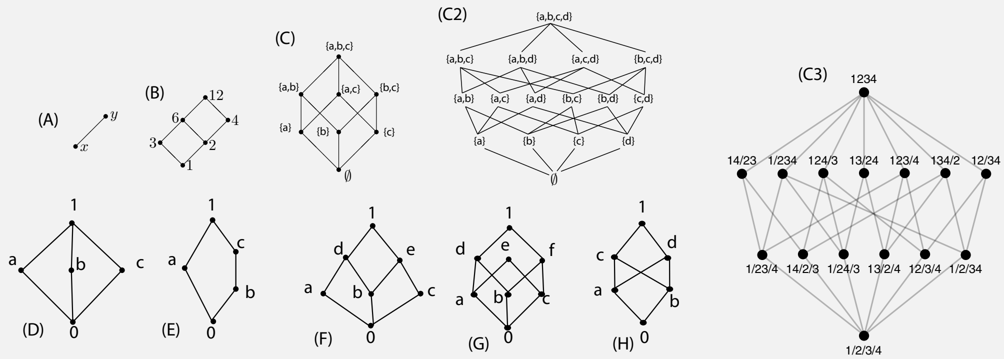
Example 14.3.3

The set $U = 2^E$ for some set E is a poset ordered by set inclusion. See Figure (C). Note that this U is self-dual.

Example 14.3.4

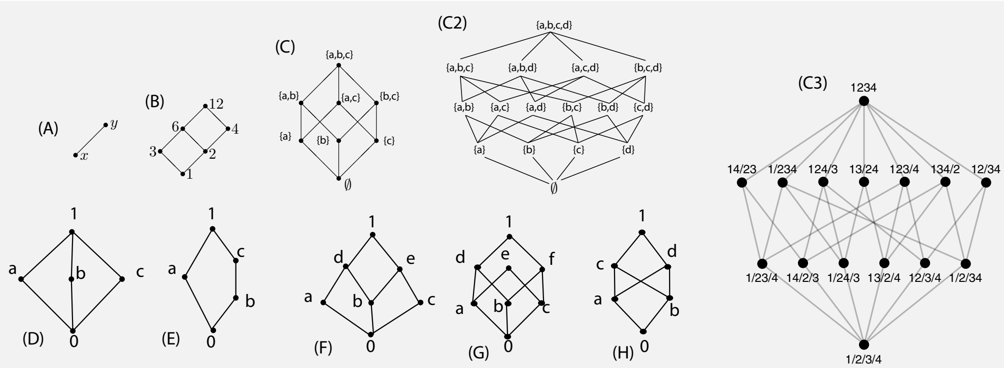
Given an n-dimensional linear (Euclidean) space $\mathbb{R}^n$. A subset of $M \subseteq \mathbb{R}^n$ is an affine set if $(1 - \lambda)x + \lambda y \in M$ whenever $x, y \in M$ and $\lambda \in \mathbb{R}$. A *linear subspace* of $\mathbb{R}^n$ is an affine set that contains the origin. Subspaces can be obtained via some A, b such that for every $y \in M$, $y = Ax + b$ for some $x \in \mathbb{R}^n$.
The set of all linear subspaces of $\mathbb{R}^n$ is a poset (ordered by inclusion), and such a set is self-dual.

Partially ordered set



- **cover** y covers x if $x \prec y$ and there exists no z such that $x \prec z \prec y$. Note that the inequalities are strict here. We write $x \sqsubset y$ if y covers x.
- A chain is saturated if it is a chain of the form $x_1 \prec x_2 \prec \cdots \prec x_n$ such that $x_1 \sqsubset x_2 \sqsubset \cdots \sqsubset x_n$ (i.e., we have a sequence of coverage relationships where x_{i+1} covers x_i for each $i < n$).
- Hasse diagram: We can draw a poset using a graph where each $x \in V$ is a node, and if $x \sqsubset y$ we draw y directly above x with a connecting edge, but no other edges.

Partially ordered set

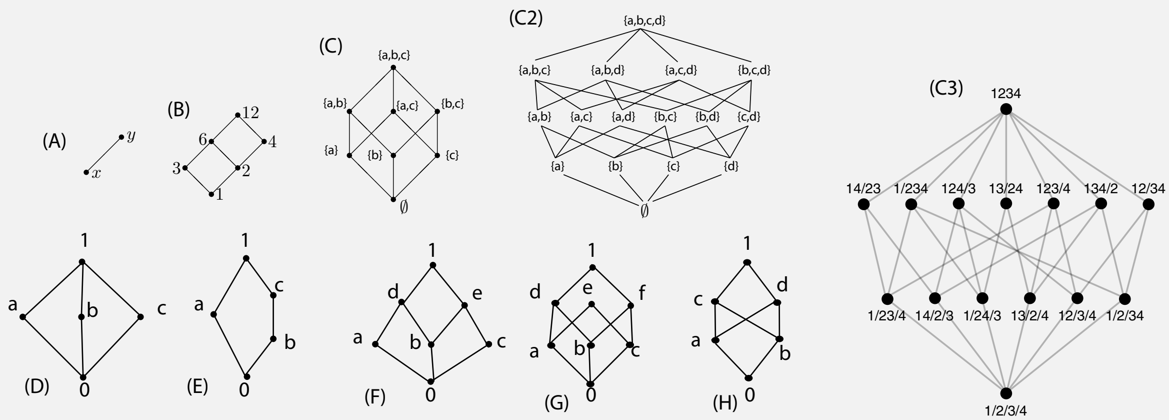


- Hasse-diagram: We can draw a poset using a graph where each $x \in V$ is a node, and if $x \sqsubset y$ we draw y directly above x with a connecting edge, but no other edges.

Theorem 14.3.5

Every non-empty finite subset $X \subseteq V$ has a minimal (and maximal) element.

Partially ordered set/Hasse diagrams examples



For example, in example (A), we see that $x \sqsubset y$. In example (B) we have $3 \preceq 12$ and $3 \sqsubset 6$ but not $3 \sqsubset 12$. Hasse diagram for dual order is obtained by turning Hasse diagram upside down.

Partially ordered set

- **least element**: any subset $X \subseteq V$, the least element of X is an element $x \in X$ such that $x \preceq y$ for all $y \in X$. The greatest element is defined similarly.
- A **minimal element** of a subset $X \subseteq V$ is an element $x \in X$ such that there exists no $y \in X$ such that $y \prec x$.

Theorem 14.3.6

Every non-empty finite subset $X \subseteq V$ has a minimal (and maximal) element.

Partially ordered set

Theorem 14.3.7

Every non-empty finite subset $X \subseteq V$ has a minimal (and maximal) element.

Proof.

Let $X = \{x_1, \ldots, x_n\}$. Define $m_1 = x_1$ and

$$m_k = \begin{cases} x_k & \text{if } x_k \prec m_{k-1} \\ m_{k-1} & \text{otherwise} \end{cases} \tag{14.1}$$

Then we have constructed $m_n \preceq m_{n-1} \preceq \cdots \preceq m_1$ meaning there is no m_k for $k < n$ such that $m_k \prec m_n$. By construction, we also have that there is no $x \in X$ with $x \prec m_n$, thus m_n is minimal. Analogously, X has a maximal element. □

Partially ordered set

Theorem 14.3.8

Every non-empty finite subset $X \subseteq V$ has a minimal (and maximal) element.

Proof.

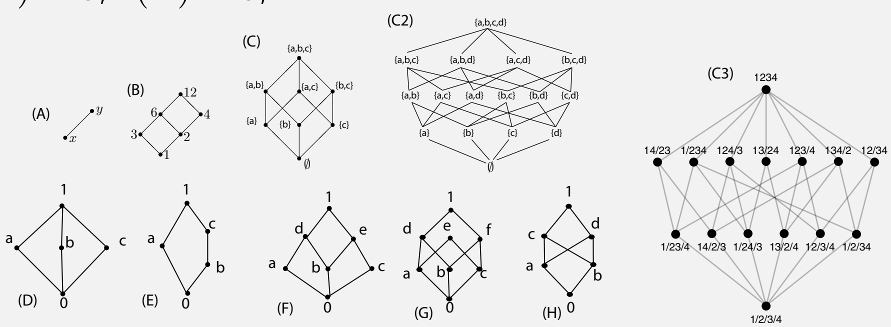
Let $X = \{x1, \ldots, x_n\}$. Define $m_1 = x_1$ and

$$m_k = \begin{cases} x_k & \text{if } x_k \prec m_{k-1} \\ m_{k-1} & \text{otherwise} \end{cases} \tag{14.2}$$

Then we have constructed $m_n \preceq m_{n-1} \preceq \cdots \preceq m_1$ meaning there is no m_k for $k < n$ such that $m_k \prec m_n$. Let $M = \{m_1, \ldots, m_n\}$. By construction, we also have that there is no $x \in X$ with $x \prec m_n$, thus m_n is minimal. □

Partially ordered set

- In chains elements, minimal equals minimum, and maximal equals maximum.
- Given a poset V, the length $\ell(V)$ is defined to be the l.u.b. of the lengths of any chains in V. That is, $\ell(V)$ is the least upper bound, i.e., smallest number not less than any chain length in V. This is finite for finite posets.
- For example, $\ell(A) = 1$, $\ell(B) = 3$, $\ell(C) = 3$, $\ell(C2) = 4$, $\ell(D) = 2$, $\ell(E) = 3$, $\ell(F) = 3$, and so on.



Partially ordered set

- The height or *dimension* of an element $x \in V$, or $l = h(x)$ is the l.u.b. of the lengths of the chains $0 = x_0 \prec x_1 \prec \dots x_l = x$ between 0 and x. Note that $h(1) = \ell(V)$ when they exist.
- $h(x) = 1$ iff $0 \sqsubset x$ and such elements (with unit height) are called "atoms" or "points" or "(ground) elements".
- graded posets. Posets might be able to be "graded" by a function $g : V \to \mathbb{Z}$ in the following way

$$x \prec y \Rightarrow g(x) < g(y) \quad \text{(G1)}$$
$$x \sqsubset y \Rightarrow g(x) + 1 = g(y) \quad \text{(G2)}$$

- A maximal chain is a chain of unique elements between two elements that can not be made any longer

Partially ordered set

- Given a poset V, the length $\ell(V)$ is defined to be the l.u.b. of the lengths of any chains in V. That is, $\ell(V)$ is the least upper bound, i.e., smallest number not less than any chain length in V.
- The height or *dimension* of an element $x \in V$, or $l = h(x)$ is the l.u.b. of the lengths of the chains $0 = x_0 \prec x_1 \prec \dots x_l = x$ between 0 and x. Note that $h(1) = \ell(V)$ when they exist. $h(x) = 1$ iff $0 \sqsubset x$ and such elements (with unit height) are called "atoms" or "points" or "(ground) elements".

Partially ordered set

- Given two points $x, y \in V$ with $x \succ y$, there might be no or multiple chains between x and y. The chains might have different lengths. There might be multiple chains that have the same maximal length.

Definition 14.3.9 (Jordan-Dedekind Chain Condition)

(or JDCC) All maximal length chains between the same endpoints have the same finite length.

Theorem 14.3.10

Let V be a poset with $0 \in V$ and where all chains are finite. Then V satisfies JDCC iff it is graded by $h(x)$ (the height function).

Proof.

Grading by $h(x)$ makes JDCC true since the length of any chain between $a \succ b$ is $h(a) - h(b)$. Conversely, given JDCC, and given $h(x)$ as defined (length of the maximal length chain from 0 to x), then G1 and G2 follow with $h(\cdot)$ as the grading function. □

Partially ordered set

Definition 14.3.11 (Jordan-Dedekind Chain Condition)

(or JDCC) All maximal length chains between the same endpoints have the same finite length.

Theorem 14.3.12

Let V be a poset with $0 \in V$ and where all chains are finite. Then V satisfies JDCC iff it is graded by $h(x)$ (the height function).

- With JDCC, if $x \sqsubset y$ then $h(x) + 1 = h(y)$.

Partially ordered set

- With JDCC, if $x \sqsubset y$ then $h(x) + 1 = h(y)$.
- When all maximal length chains between the same endpoints have the same finite length, then we say that the poset is graded by the height. In this case, we say that element x has height or rank $h(x)$.
- The height (rank) function in this case is unique. If $x \preceq y$ then $\ell(x, y) = h(y) - h(x)$ is the length between x and y.
- In fact, any graded poset satisfies JDCC, and hence is graded by the height function. Hence, it sufficies for poset to be graded if it is graded by the height function.

Partially ordered set

- With JDCC, element x has height or *rank* $h(x)$. The height (rank) function in this case is unique. If $x \preceq y$ then $\ell(x, y) = h(y) - h(x)$ is the length between x and y.
- We say a poset is "graded" if it is graded by the height function.

Lattices

- Birkoff lattices are an important algebraic structure to understand when studying submodular functions.
- Early work on lattices occurred around 1990 by Richard Dedekind, on what were called "modular" lattices (we'll see this below).
- Birkhoff concentrated on the mathematics of lattices in the 20th century.
- Lattices are a useful algebraic structure, but also shed some light on one of the origins of submodular functions (including their name).

Lattice defined

- Given $X \subset V$, $y \in V$ is an upper bound of X if $x \preceq y$ for all $x \in X$. Note that y need not be in X.
- If y is a least upper bound (l.u.b. X or just sup$X$), then $y \preceq z$ for any other upper bound z.
- The l.u.b. if it exists is unique since if y_1 and y_2 are both l.u.b.'s then $y_1 \preceq y_2$ and $y_2 \preceq y_1$, or $y_1 = y_2$.
- Dual definitions for lower bound and greatest lower bound (g.l.b.).

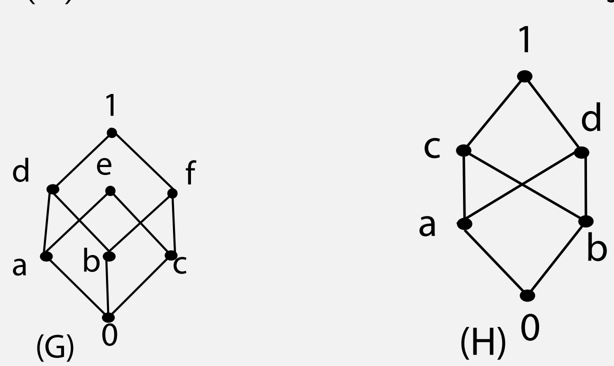
Definition 14.4.1 (lattice)

A *lattice* is a poset V such that any two elements $x, y \in v$ have a g.l.b. or meet denoted by $x \wedge y \in V$, and also have a l.u.b. or join denoted by $x \vee y \in V$. A lattice is complete when all subsets $X \subseteq V$ have both a l.u.b. and a g.l.b. (note that join and meet is defined on pairs, but l.u.b. and g.l.b. can be defined on any subset of V, even of size 1).

- Note again, that such l.u.b.'s and g.l.b.'s are unique if they exist.

Lattices

- Any finite lattice or lattice of finite length is complete. Note that the reverse need not hold (a complete lattice need not be finite). The reals are not complete but the extended reals are complete. The rationals are not complete (but the rationals extended with a $\pm\infty$ is complete). 2^E for some set E is complete. Note that E can be countably or uncountably infinite.
- All of the figures above are lattices except for (G) and (H). (G) is not a lattice since for example $e \vee f$ does not exist, nor does any join with e exist. (H) is not a lattice since there are two joins for a and b.



- Any non-empty lattice contains a greatest element $1 \in V$ and a least element $0 \in V$.

Lattices

Definition 14.4.2 (sublattice)

A sublattice of a lattice is a subset $X \subseteq V$ such that join and meet are closed within X (for all $x, y \in X$, $x \vee y \in X$ and $x \wedge y \in X$). A sublattice is a lattice.

- Given any $x \preceq y$, then all elements $\{z : x \preceq z \preceq y\}$ form a sublattice. We note that in such case, we say that $[x, y]$ form a (closed) interval in the lattice, and we have that the (closed) interval $[x, y]$ of all elements $z \in L$ such that $x \preceq z \preceq y$ is a sublattice.
- A convex subset X of a poset V is a subset such that for all $x, y \in V$ with $x \preceq y$, $\{z : x \preceq z \preceq y\} \subseteq X$. A subset X of a *lattice* V is a convex sublattice if $x, y \in X$ imply that $\{z : x \wedge y \preceq z \preceq x \vee y\} \subseteq X$.
- Obviously, 2^E for some set E is a lattice, with join/meet being union/intersection. See Figure(C).

Lattices

Example 14.4.3

Given lattices U, V, we can form the direct product UV by forming pairs $\{(u, v) : u \in U, v \in V\}$ and ordered so that $(u_1, v_1) \preceq (u_2, v_2)$ iff $u_1 \preceq u_2$ in U and $v_1 \preceq v_2$ in V. The direct product of two lattices is a lattice.

Theorem 14.4.4

In any poset V, the operations of meet and join satisfy the following laws, whenever the associated expressions exist.

$$x \wedge x = x, x \vee x = x \qquad \textit{(Idempotent)} \qquad \text{(L1)}$$

$$x \wedge y = y \wedge x, x \vee y = y \vee x \qquad \textit{(Commutative)} \qquad \text{(L2)}$$

$$x \wedge (y \wedge z) = (x \wedge y) \wedge z, x \vee (y \vee z) = (x \vee y) \vee z \qquad \textit{(Associative)} \qquad \text{(L3)}$$

$$x \wedge (x \vee y) = x \vee (x \wedge y) = x \qquad \textit{(Absorption)} \qquad \text{(L4)}$$

$$x \preceq y \iff x \wedge y = x \textit{ and } x \vee y = y \qquad \textit{(Consistency)} \qquad \text{(CON)}$$

Note the above works for posets, not necessary for it to be a lattice.

Lattices

Theorem 14.4.5

Given a poset V with $0 \in V$, then for all $x \in V$,

$$0 \wedge x = 0 \text{ and } 0 \vee x = x \tag{14.3}$$

Theorem 14.4.6

In any lattice, the operations of join and meet are order-preserving in the following sense:

$$y \preceq z \Rightarrow x \wedge y \preceq x \wedge z \text{ and } x \vee y \preceq x \vee z \tag{14.4}$$

Distributive Inequalities

Theorem 14.4.7

In any lattice, the following distributive inequalities hold for all $x, y, z \in V$:

$$x \wedge (y \vee z) \succeq (x \wedge y) \vee (x \wedge z) \tag{14.5a}$$
$$x \vee (y \wedge z) \preceq (x \vee y) \wedge (x \vee z) \tag{14.5b}$$

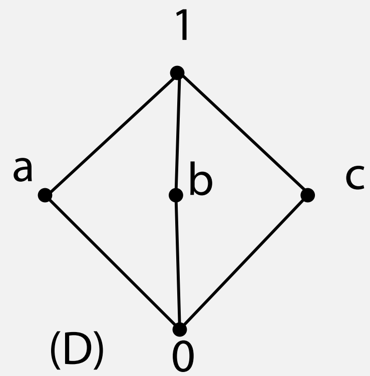
Proof.

We have $x \wedge y \preceq x$ and $x \wedge y \preceq y \preceq y \vee z$. Therefore, $x \wedge y \preceq x \wedge (y \vee z)$. Also, since $x \wedge z \preceq x$ and $x \wedge z \preceq z \preceq y \vee z$, we have $x \wedge z \preceq x \wedge (y \vee z)$. Thus, $x \wedge (y \vee z)$ is an upper bound of both $x \wedge y$ and $x \wedge z$, which means that it is an upper bound of their join.
Eq 14.5b is by duality. □

- Note, this does not mean the lattice is distributive, which we define in a few slides below.

Distributive Inequalities

- Note that these are inequalities and they hold in any lattice. Equality might not hold for all lattices, consider figures (D) and (E).

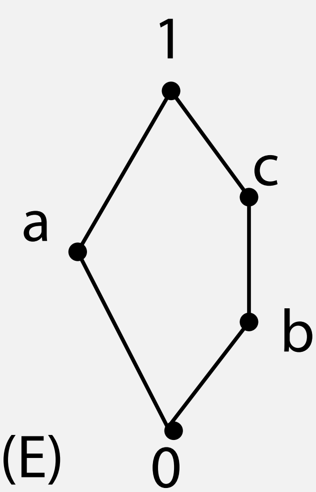


(D)

- For example, in (D), we have that $a \wedge (b \vee c) = a \wedge 1 = a$ but $(a \wedge b) \vee (a \wedge c) = 0 \vee 0 = 0$ and obviously $a \succ 0$.
- Also, in (D) we have $a \vee (b \wedge c) = a \vee 0 = a \prec (a \vee b) \wedge (a \vee c) = 1 \wedge 1 = 1$.

Distributive Inequalities

- Note that these are inequalities and they hold in any lattice. Equality might not hold for all lattices, consider figures (D) and (E).



(E)

- In (E), we have that $c \wedge (a \vee b) = c \wedge 1 = c \succ (c \wedge a) \vee (c \wedge b) = 0 \vee b = b$
- Also, in (E), we have $b \vee (a \wedge c) b \vee 0 = b \prec (b \vee a) \wedge (b \vee c) = 1 \wedge c = c$

Modular inequality

Theorem 14.4.8

In any lattice, the following modular inequalities holds for all $x, y, z \in V$:

$$x \preceq z \Rightarrow x \vee (y \wedge z) \preceq (x \vee y) \wedge z \tag{14.6}$$

Proof.

We have that $x \preceq x \vee y$ and are given $x \preceq z$. Then, $x \preceq (x \vee y) \wedge z$. Also, $y \wedge z \preceq y \preceq x \vee y$ and with $y \wedge z \preceq z$ gives $y \wedge z \preceq (x \vee y) \wedge z$. Then $x \vee (y \wedge z) \preceq (x \vee y) \wedge z$. □

- The term "modular" somehow comes from abstract algebra, where a R-module is an abstract system that generalizes $(\mathbb{R}, \mathbb{R}^n)$ (i.e., a vector field with scalar multiplication). An R-module ends up being a lattice that satisfies this property with equality (see below).

Distributive Lattices

- A lattice is distributive if the aforementioned distributive inequality is an equality. Note that as mentioned above, the distributive inequality holds for all lattices, but not with equality.
- Some lattices are such that the distributive inequality is an equality everywhere, and these are called distributive lattices. Only one quality is necessary since:

Theorem 14.5.1

In any lattice, the following are equivalent:

$$x \wedge (y \vee z) = (x \wedge y) \vee (x \wedge z) \qquad \forall x, y, z \tag{14.7a}$$

$$x \vee (y \wedge z) = (x \vee y) \wedge (x \vee z) \qquad \forall x, y, z \tag{14.7b}$$

It is important to note the $\forall x, y, z$ since this is not true only for individual elements. Note moreover that this means that the operators $\vee = +$ and $\wedge = \cdot$ do not form a lattice over $\mathbb{R}$.

Distributive Lattices

Theorem 14.5.2

In any lattice, the following are equivalent:

$$x \wedge (y \vee z) = (x \wedge y) \vee (x \wedge z) \qquad \forall x, y, z \tag{14.8a}$$
$$x \vee (y \wedge z) = (x \vee y) \wedge (x \vee z) \qquad \forall x, y, z \tag{14.8b}$$

Proof.

Take as given the 2nd equation and show the first. Then

$$(x \wedge y) \vee (x \wedge z) = [(x \wedge y) \vee x] \wedge [(x \wedge y) \vee z] \quad \text{by the 2nd eq} \tag{14.9}$$
$$= x \wedge [(x \wedge y) \vee z] \quad x \wedge y \preceq x \tag{14.10}$$
$$= x \wedge [(x \vee z) \wedge (y \vee z)] \quad \text{by the 2nd eq} \tag{14.11}$$
$$= x \wedge (x \vee z) \wedge (y \vee z) \quad \text{associative} \tag{14.12}$$
$$= x \wedge (y \vee z) \quad x \vee z \succeq x \tag{14.13}$$

Distributive Lattices

- Note that any chain is a distributive lattice. The dual of any distributive lattice is distributive.
- Thus a lattice is distributive if either of the above equalities hold.

Example 14.5.3

Let $V = \mathbb{Z}^+$ be the set of positive integers and let $x \preceq y$ mean that x divides y. I.e., $2 \preceq 4$ but $2 \not\preceq 5$. Then this is lattice with $x \vee y =$ l.c.m.(x, y) and $x \wedge y =$ g.c.d.(x, y). It is also distributive. Again consider figure (B).

Theorem 14.5.4 (identity)

In a distributive lattice, if $z \wedge x = z \wedge y$ and $z \vee x = z \vee y$ then $x = y$.

Modular Lattices

- In Theorem 14.4.8, we also defined the modular inequality. We can strengthen this as well to get what is known as the modular identity.

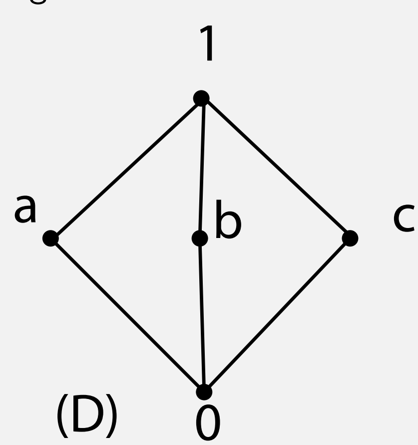
Definition 14.6.1 (modular identity)

$$\forall x, y, z, \quad \text{If } x \preceq z, \text{ then } x \vee (y \wedge z) = (x \vee y) \wedge z. \tag{L5}$$

- Clearly any distributive lattice satisfies the modular identity since when $x \preceq z$ we have that $x \vee z = z$ and from the 2nd of the distributive lattice equalities (i.e., $x \vee (y \wedge z) = (x \vee y) \wedge (x \vee z)$) we get the modular identity.
- Easy way to remember. x,y,z and $x \vee \big(y \wedge z\big) = \big(x \vee y\big) \wedge z$
- The term "modular" comes from abstract algebra, where a R-module is an abstract system that generalizes $(\mathbb{R}, \mathbb{R}^n)$ (i.e., a vector field with scalar multiplication). An R-module ends up being a lattice that satisfies this identity.

Modular Lattices

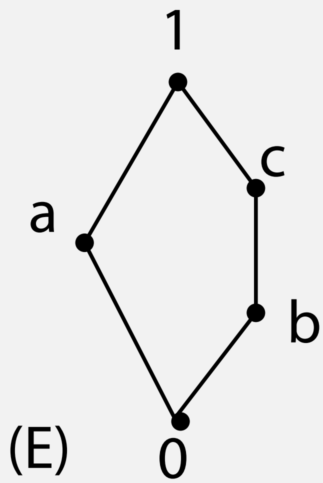
- Not every lattice is modular. Figure (D) is modular but not distributive. We already saw that (D) is not distributive since it is strict for certain assignments. It is modular though.



(D)

Modular Lattices

- Figure (E) is neither modular nor distributive. We saw that it was not distributive since it achieved strictness in the distributive inequalities. It is not modular since: take $b \preceq c$, then $b \vee (a \wedge c) = b \vee 0 = b \prec (b \vee a) \wedge c = 1 \wedge c = c$, so modular equality is violated.



Theorem 14.6.2

Any non-modular lattice V contains the lattice in Figure (E) as a sublattice.

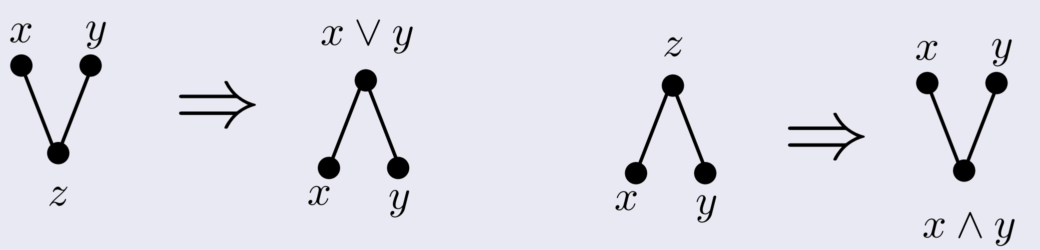
Modular Lattices

Theorem 14.6.3

A necessary and sufficient condition for a modular lattice is to have both:

A) Upper-Semimodularity if x and y cover z and $x \neq y$ then $x \vee y$ covers both x and y, and

B) Lower-Semimodularity if z covers x and y and $x \neq y$ then x and y both covers $x \wedge y$.



- Thus, upper-semimodularity means that if $z \sqsubset x$ and $z \sqsubset y$, and if $x \neq y$, then $x \sqsubset (x \vee y)$ and $y \sqsubset (x \vee y)$.
- Thus, lower-semimodularity means that if $x \sqsubset z$ and $y \sqsubset z$, and if $x \neq y$, then $(x \wedge y) \sqsubset x$ and $(x \wedge y) \sqsubset y$.

Upper-Semimodularity vs. Lower-Semimodularity

- As we will see, the first (left, upper-semimodularity) equation implies submodularity on the dimension (height function)
- The second (right, upper-semimodularity) equation implies supermodularity on the dimension (height) function.
- Both together imply modularity on the dimension function.

Semi-modular/Submodular Lattices

Theorem 14.7.1

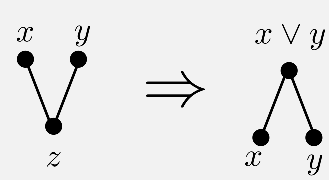
Let L be a finite lattice. The following two conditions are equivalent:

(i) *L is graded, and the height function $h(\cdot)$ of L satisfies the (what we know as the submodular) inequality for all $x, y \in L$.*

$$h(x) + h(y) \geq h(x \vee y) + h(x \wedge y) \tag{14.14}$$

(ii) *If x and y both cover z, then $x \vee y$ covers both x and y*

- Condition (ii) is visualized as:



Semi-modular/Submodular Lattices: (i) $\Rightarrow$(ii)

h submodular $\Rightarrow \left\{(z \sqsubset x, z \sqsubset y) \Rightarrow (x \sqsubset (x \vee y)), (y \sqsubset (x \vee y))\right\}$.

- Suppose x and y cover z.
- Note that if x and y cover z then since L is a lattice, $z = x \wedge y$.
- Then we have $h(x) = h(y) = h(x \wedge y) + 1$.
- Also, since x and y are distinct, and since they both cover z we can't have (w.l.o.g.) $x \preceq y$, and thus $h(x \vee y) > h(x) = h(y)$.
- Hence by (i), we have

$$h(x) + h(y) - h(x \wedge y) \geq h(x \vee y) > h(x \wedge y) + 1 \tag{14.15}$$

- or

$$h(x \wedge y) + 2 \geq h(x \vee y) > h(x \wedge y) + 1 \tag{14.16}$$

- giving $h(x \vee y) = h(x \wedge y) + 2 = h(x) + 1 = h(y) + 1$, so that $x \vee y$ covers both x and y.

...

Semi-modular/Submodular Lattices: (ii) $\Rightarrow$ (i)

$\left\{(z \sqsubset x, z \sqsubset y) \Rightarrow (x \sqsubset (x \vee y)), (y \sqsubset (x \vee y))\right\} \Rightarrow h$ submodular.

- Suppose L is not graded, and let $[u, v]$ be an interval of L of minimal length that is not graded (so all smaller length intervals are graded).
- Then there are elements x_1, x_2 of $[u, v]$ where each of x_1 and x_2 cover u,
- By the minimality, all maximal chains of each interval $[x_i, v]$ have the same length ℓ_i where $\ell_1 \neq \ell_2$.
- By (ii), there are saturated chains in $[x_i, v]$ of the form $x_i \prec x_1 \vee x_2 \prec y_1 \prec y_2 \prec \cdots \prec y_k = v$, contradicting $\ell_1 \neq \ell_2$.
- Hence L is graded (i.e., every maximal chain has the same length, i.e., JDCC holds).

...

Semi-modular/Submodular Lattices: (ii) $\Rightarrow$ (i)

$\Big\{(z \sqsubset x, z \sqsubset y) \Rightarrow (x \sqsubset (x \vee y)), (y \sqsubset (x \vee y))\Big\} \Rightarrow h$ submodular.

- Now suppose there is a pair $x, y \in L$ violating the submodularity inequality, i.e., with

$$h(x) + h(y) < h(x \vee y) + h(x \wedge y) \tag{14.17}$$

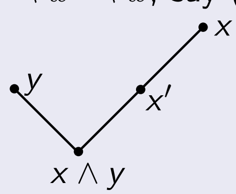
  and choose such a pair first with $\ell(x \wedge y, x \vee y)$ minimal, and then (second) with $h(x) + h(y)$ minimal.
- By (ii), we cannot have both x and y covering $x \wedge y$ (because if we did, then $h(x) = h(x \wedge y) + 1$, $h(y) = h(x \wedge y) + 1$, and (ii) gives that $h(x \vee y) = h(x) + 1 = h(y) + 1$, and we would have the submodular inequality at equality).

...

Semi-modular/Submodular Lattices: (ii) $\Rightarrow$ (i)

$\Big\{(z \sqsubset x, z \sqsubset y) \Rightarrow (x \sqsubset (x \vee y)), (y \sqsubset (x \vee y))\Big\} \Rightarrow h$ submodular.

- Thus assume that $x \wedge y \prec x' \prec x$, say (w.l.o.g.)



.

- By the minimality of $\ell(x \wedge y, x \vee y)$ and $h(x) + h(y)$, we have

$$h(x') + h(y) \geq h(x' \wedge y) + h(x' \vee y). \tag{14.18}$$

...

Semi-modular/Submodular Lattices: (ii) $\Rightarrow$ (i)

$\left\{(z \sqsubset x, z \sqsubset y) \Rightarrow (x \sqsubset (x \vee y)), (y \sqsubset (x \vee y))\right\} \Rightarrow h$ submodular.

- Now $x' \wedge y = x \wedge y$, so Eq. 14.17 and Eq. 14.18 together imply that

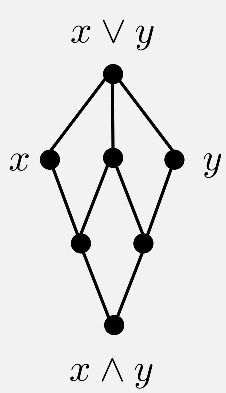
$$h(x) + h(x' \vee y) < h(x') + h(x \vee y). \tag{14.19}$$

- Since $x \succ x'$, we have $x \vee (x' \vee y) = (x \vee x') \vee y = x \vee y$.
- Also, by the modular inequalities (with $x \leftarrow x', y \leftarrow y, z \leftarrow x$), we have $x \wedge (x' \vee y) \succeq x' \vee (y \wedge x) \succeq x'$.
- Hence setting $X = x$, $Y = x' \vee y$. This gives $X \vee Y = x \vee y$, and $X \wedge Y \succeq x' \succ x$.
- Thus, we have found a pair $X, Y \in L$ with $h(X) + h(Y) < h(X \wedge Y) + h(X \vee Y)$ and a strictly shorter length $\ell(X \wedge Y, X \vee Y) < \ell(x \wedge y, x \vee y)$,
- This contradicts the minimality of $\ell(x \wedge y, x \vee y)$.
- The proof is complete.

...

Submodular Lattices

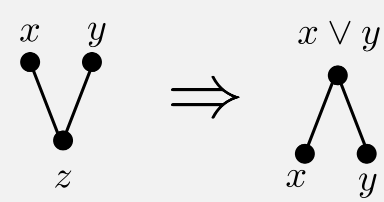
The next figure is an example of an upper-semimodular (or a "submodular") lattice over 7 elements.



height
3
2
1
0

submodularity
h(x)+h(y)
> h($x \vee y$)
+ h($x \wedge y$)
2 + 2 > 3 + 0

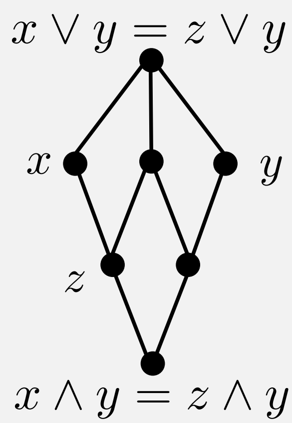
- Can verify that for all x, y, z that:



- Hence, the lattice is upper-semimodular (submodular), as the height function is submodular on the lattice.

Submodular Lattices

The next figure is an example of an upper-semimodular (or a "submodular") lattice over 7 elements.



- Note, the lattice is not modular since it violates the modular equality ($\forall x, y, z, \quad z \preceq x \Rightarrow (z \vee (y \wedge x) = (z \vee y) \wedge x)$).
- I.e., on the left, $z \vee (y \wedge x) = z$ but $(z \vee y) \wedge x = x$
- Flip the Hasse diagram up-side down to get a lower-semimodular (or "supermodular") lattice and corresponding height function.

Ideal in a Lattice

Definition 14.8.1 (ideal)

An ideal is a nonvoid subset J of a lattice L with the properties

$$\forall a \in J, x \in L, \quad x \preceq a \Rightarrow x \in J \tag{14.20}$$

$$\forall a \in J, \quad b \in J \Rightarrow a \vee b \in J. \tag{14.21}$$

The dual concept (in a lattice) is called a dual ideal (or a meet ideal).

Proposition 14.8.2

J is an ideal when $a \vee b \in J$ iff $a \in J, b \in J$ (closure under join).

Example 14.8.3

In 2^E, take any $A \subseteq E$, then $L(A) = \{B : B \subseteq A\}$ is an ideal in a set lattice.

Ideal in a Lattice

Definition 14.8.4

Given an element $a \in L$ in a lattice, the set $L(a)$ of all elements $\{x : x \preceq a, x \in L\}$ is an ideal, and is called a principle ideal.

In fact, in any finite lattice, every (nonvoid) ideal is a principle ideal. In fact, we have:

Theorem 14.8.5

The set of all ideals'of any lattice L, ordered by inclusion, itself forms a lattice. The set of all principal ideals in L forms a sublattice of this lattice, which is isomorphic with L.

Example 14.8.6

Consider 2^E. Then for any $A \subseteq E$, we see that $L(A) = \{B : B \subseteq A\}$ is an ideal. Also, we can see that the set of sets $\{L(A) : A \subseteq E\}$ is isomorphic to 2^E and also forms a lattice.

Complement and Complemented Lattices

Definition 14.8.7

A lattice with a 0 and 1 is complemented if for all $x \in L$ there exists a $y \in L$ such that $x \vee y = 1$ and $x \wedge y = 0$. A lattice is relatively complemented if every interval $[x, y]$ is complemented (w.r.t. the interval, with x taking the role of 0 and y taking the role of 1).

Recall, an atom of a finite lattice is an element covering 0.

Proposition 14.8.8

Any complemented modular lattice is relatively complemented.

Proposition 14.8.9

In a complemented modular lattice of finite length, every element is the join of those elements which it contains.

Boolean Lattices

Definition 14.8.10

A Boolean lattice is a complemented distributive lattice.

Theorem 14.8.11

In any Boolean lattice, each element x has a unique complement x'. Moreover, we have

$$x \wedge x' = 0, \qquad x \vee x' = 1 \tag{L1}$$

$$(x')' = x, \tag{L2}$$

$$(x \wedge y)' = x' \vee y', \qquad (x \vee y)' = x' \wedge y' \tag{L3}$$

Join Irreducible

Definition 14.8.12

An element x of a lattice is called join irreducible if $y \vee z = x$ implies $y = x$ or $z = x$ (ie, if x is the join of two elements, it must be one of those elements). Equivalently, an element x of a lattice is join irreducible if one cannot write $x = y \vee z$ where $y \prec x$ and $z \prec x$ (ie, x is not the join of two strictly smaller elements).

Proposition 14.8.13

If all chains in a lattice are finite, then every $a \in L$ can be represented as a join $a = x_1 \vee \ldots \vee x_n$ of a finite number of join irreducible elements.

Proposition 14.8.14

In any complemented modular lattice, all join irreducible elements are atoms.

Ring of sets

Definition 14.8.15 (ring family)

A ring of sets (or ring family) is a family Φ of subsets of a set E which contains with any two sets S and T also their (set-theoretic) intersection $S \cap T$ and union $S \cup T$. A field of sets is a ring of sets which contains with any S also its set complement $E \setminus S$'

- Thus, any ring of sets under the natural ordering $S \subset T$ forms a distributive lattice.

Join irreducible, ground elements, Boolean lattices

Theorem 14.8.16

Let L be any distributive lattice of length n. Then the poset X of join-irreducible elements $x \succ 0$ has order n and, moreover, $L \simeq \mathbf{2}^X$

- The join-irreducible elements of a distributive lattice constitute a form of set of "ground elements" which generate the distributive lattice.
- Thus, any distributive lattice of length n is isomorphic with a ring of subsets of a set E of n elements.
- The next result is perhaps not so surprising.

Theorem 14.8.17

Every Boolean lattice of finite length n is isomorphic with the field of all subsets of a set of $|E| = n$ elements, namely 2^E.

Sources for Today's Lecture

- Birkhoff, "Lattice Theory", 1967.