Introduction to Möbius functions.

Bogdan Chornomaz

September 25, 2020

Incidence algebras.

For a finite lattice L (locally finite poset P), an **incidence algebra** of L is a set of functions $\{\alpha : I(L) \to \mathbb{Z}\}$, where $I(L) = \{(x, y) \in L^2 \mid x \leq y\}$ with an associative **convolution**:

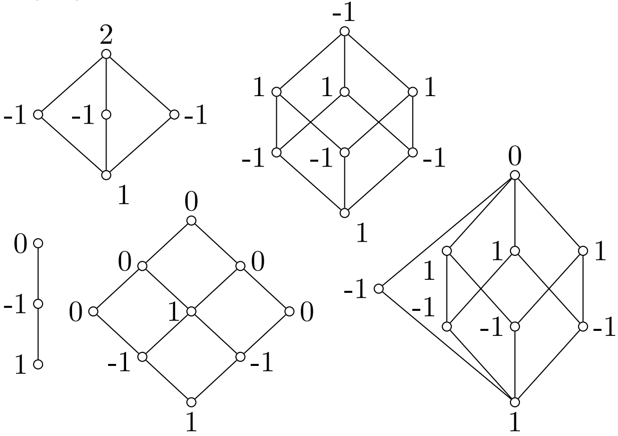
$$\alpha * \beta(x, y) = \sum_{x \leq z \leq y} \alpha(x, z)\beta(z, y).$$

Several special elements in incidence algebra are:

- $\delta(x, y) = \begin{cases} 1, x = y, \\ 0, x < y; \end{cases}$ δ is a unit of the algebra;
- $\zeta \equiv 1$;
- μ - unique left and right inverse of ζ, i.e. $\mu * \zeta = \delta$, $\zeta * \mu = \delta$;
- $\mu(x, y) = \begin{cases} 1 & x = y; \\ -\sum_{x \leq z < y} \mu(x, z) & \textit{otherwise}. \end{cases}$

Some examples

$\mu(0, \cdot)$



Möbius function of the boolean lattice is $(-1)^l$, where l is the length of the interval.

It is zero above any non-atomic join-irreducible element...

but it also can be zero in other cases.

Why right inverse?

The explicit formula for μ comes from expanding the formula $\mu * \zeta = \delta$:

$$\begin{cases} \mu * \zeta(x,x) = \delta(x,x) & \Leftrightarrow \qquad \mu(x,x) = 1; \\ \mu * \zeta(x,y) = \delta(x,y) & \Leftrightarrow \qquad \sum_{x \le z \le y} \mu(x,z) = 0. \end{cases}$$

Hence $\mu(x,y) = -\sum_{x \le z < y} \mu(x,z)$ for $x < y$ and this enables to define μ iteratively in a unique fashion. But saying that μ is the right inverse of ζ, $\zeta * \mu = \delta$, is equivalent to $\mu(x,y) = -\sum_{x < z \le y} \mu(z,y)$. How does that follow?

First note, that any α for which $\alpha(x,x) \equiv 1$ has left and right inverse. Let θ be a left inverse for μ, $\theta * \mu = \delta$, then

$$\begin{aligned} \zeta * \mu &= (\theta * \mu) * (\zeta * \mu) \\ &= \theta * (\mu * \zeta) * \mu = \theta * \mu = \delta. \end{aligned}$$

In fact, matrices

Let $n = |L|$ and let us enumerate the elements of L in any compatible order (that is, $x_i \leq_L x_j$ implies $i \leq j$). Now, for α from the incidence algebra of L we put into a correspondence an integer $n \times n$ matrix M_α, defined as:

$$M_\alpha(i,j) = \begin{cases} \alpha(x_i, x_j) & x_i \leq_L x_j \\ 0 & \text{otherwise} \end{cases}$$

then

$$\begin{aligned} M_\alpha M_\beta(i,j) &= \sum_k M_\alpha(i,k) M_\beta(k,j) \\ &= \sum_{k\colon x_i \leq_L x_k \leq_L x_j} \alpha(x_i, x_k)\beta(x_k, x_j) = M_{\alpha * \beta}(i,j). \end{aligned}$$

Also, as the order is compatible, all those matrices are upper-triangular. And if $\alpha(x,x) \equiv 1$ then $\det M_\alpha = 1$ and thus M_α is invertible.

Products

Lemma

$$\mu_{L\times K}\big((x_1,x_2),(y_1,y_2)\big) = \mu_L(x_1,y_1)\mu_K(x_2,y_2).$$

Proof. By induction:

$$\begin{aligned}
\mu_{L\times K}\big((x_1,x_2),(y_1,y_2)\big) &= -\sum_{(x_1,x_2)\le(z_1,z_2)<(y_1,y_2)} \mu_{L\times K}\big((x_1,x_2),(z_1,z_2)\big) \\
&= -\sum_{\substack{x_1\le z_1\le y_1;\;\; x_2\le z_2\le y_2 \\ (z_1,z_2)\ne(y_1,y_2)}} \mu_L(x_1,z_1)\mu_K(x_2,z_2) \\
&= -\sum_{x_1\le z_1\le y_1} \mu_L(x_1,z_1) \cdot \sum_{x_2\le z_2\le y_2} \mu_K(x_2,z_2) + \mu_L(x_1,y_1)\mu_K(x_2,y_2) \\
&= \mu_L(x_1,y_1)\mu_K(x_2,y_2)
\end{aligned}$$

Möbius inversion formula

Theorem

Let $f, g\colon L \to \mathbb{Z}$. *Then*

$$f(x) = \sum_{y \leq x} g(y) \Leftrightarrow g(x) = \sum_{y \leq x} f(y)\mu(y, x).$$

Proof. We use matrix representation. If we "interpret" f and g as $1 \times n$ vectors then

$$f(x) = \sum_{y \leq x} g(y) \quad \cong \quad f = g \cdot M_\zeta$$

thus $g = g \cdot M_\delta = g \cdot M_\zeta M_\mu = f \cdot M_\mu$.

Similarly, $f(x) = \sum_{y \geq x} g(y) \;\cong\; f^t = M_\zeta \cdot g^t$, hence

Theorem

$$f(x) = \sum_{y \geq x} g(y) \Leftrightarrow g(x) = \sum_{y \geq x} \mu(x, y) f(y).$$

Inclusion-exclusion principle

Note that $B_n = \mathbf{2} \times \cdots \times \mathbf{2}$. Hence $\mu_{B_n}(S,T) = (-1)^{|T|-|S|}$, where $S, T \subseteq \mathbf{n}$. Now we can derive inclusion-exclusion principle:

Let $X = A_1 \cup \cdots \cup A_n$. Let $f(S) = \left|\bigcap_{s \in S} A_s\right|$ and $g(S) = \left|\bigcap_{s \in S} A_s \cap \bigcap_{s' \notin S} (X - A_{s'})\right|$. Then

$$f(S) = \sum_{S \subseteq T} g(T) \quad \Leftrightarrow \quad g(S) = \sum_{S \subseteq T} (-1)^{|T|-|S|} f(T)$$

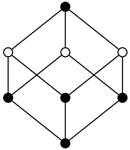
Plug $S = \emptyset$ to the second equation. Note that $g(\emptyset) = 0$ and $f(\emptyset) = |X|$

$$0 = \sum_{T} (-1)^{|T|} f(T) \quad \Leftrightarrow \quad f(\emptyset) = \sum_{T \neq \emptyset} (-1)^{|T|+1} f(T).$$

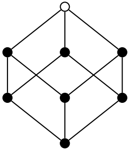
And the latter is inclusion-exclusion formula.

Can't live without shattering.

$\mathcal{F}$



$Sh(\mathcal{F})$



$\mathcal{F}$

$Sh(\mathcal{F})$

Just in case you forgot.

For a finite lattice L and $F \subseteq L$, F **shatters** an element $y \in L$, iff

$$\forall z \leq y \quad \exists x \in F \quad \text{s.t.} \quad z = y \wedge x.$$

The set of elements shattered by F is denoted $Sh(F)$, and it is an order-ideal.

L satisfies Sauer-Shelah-Perles lemma (is SSP), if for any $F \subseteq L$ it holds: $|F| \leq |Sh(F)|$.

Theorem (Babai, Frankl)

If a lattice L has a non-vanishing Möbius function μ, then it is SSP.

Proof of the theorem.

Theorem (Babai, Frankl)

If a lattice L has a non-vanishing Möbius function μ, then it is SSP.

Fix $F \subseteq L$ and consider an F-dimensional vector space $\mathbb{R}[F]$ over $\mathbb{R}$ (doesn't really matter over what). That is, elements of $\mathbb{R}[F]$ are functions $f\colon F \to \mathbb{R}$.

For $x \in L$ let $\chi_x \in \mathbb{R}[F]$ be defined as $\chi_x = \mathbf{1}_{\{y \in F \mid y \geq x\}}$. That is,

$$\begin{cases} 1 & y \geq x; \\ 0 & \text{otherwise.} \end{cases}$$

The set $\{\chi_x \mid x \in F\}$ is a basis of $\mathbb{R}[F]$. We claim that $\{\chi_x \mid x \in Sh(F)\}$ is also a basis. The claim then follows by comparing dimensions.

Note. Functions χ_x are defined only on F, but they are defined for any $x \in L$.

Proof of the theorem.

Theorem (Babai, Frankl)

If a lattice L has a non-vanishing Möbius function μ, then it is SSP.

Claim. If x is not shattered by F through y, then

$$\chi_x = -\sum_{y \leq z < x} \frac{\mu(y,z)}{\mu(y,x)} \chi_z$$

Proof (of the claim). This is equivalent to

$$\sum_{y \leq z \leq x} \mu(y,z)\chi_z = 0. \qquad (*)$$

Now, why is this obvious?

1. Let's for a moment think of χ_z as defined on all L, not only on F.
2. The function $(*)$ is constant on partitions $P_z = \{u \mid u \wedge x = z\}$, $z \in [y,x]$.
3. By the definition of μ, $(*)$ is 0 on P_z for $z \neq y$. Only P_y is a trouble.
4. But $P_y \cap F = \emptyset$, so $(*)$ as a function on F doesn't see it. □

Proof of the theorem.

Theorem (Babai, Frankl)

If a lattice L has a non-vanishing Möbius function μ, then it is SSP.

And that's it! Once again:

1. The dimension of $\mathbb{R}[F]$ is $|F|$, $\{\chi_x \mid x \in F\}$ is its basis;
2. Using the claim, $\chi_x \in \mathsf{Lin}\{\chi_z \mid z < x\}$, whenever $x \notin Sh(F)$;
3. As $Sh(F)$ is an order-ideal, by iteratively applying the claim, we get $\chi_x \in \mathsf{Lin}\{\chi_z \mid z \in Sh(F)\}$, for any $x \in L$.
4. Thus, $\mathbb{R}[F] = \mathsf{Lin}\{\chi_x \mid x \in F\} \leq \mathsf{Lin}\{\chi_z \mid z \in Sh(F)\}$;
5. So $|F| = \dim \mathbb{R}[F] \leq \dim \mathsf{Lin}\{\chi_z \mid z \in Sh(F)\} \leq |Sh(F)|$.

Geometric lattices

Let L be a finite lattice, which is:

1. atomic;
2. *graded*, that is, there is a rank function $r\colon L \to \mathbb{Z}$ such that every maximal chain in $[0, x]$ has length $r(x)$;
3. *semimodular*, that is, $r(x) + r(y) \geq r(x \wedge y) + r(x \vee y)$.

Such lattices are called *geometric*. Why do we care?

A *matroid* $\mathcal{M}$ is a nonempty order-ideal $\mathcal{I} \subseteq 2^X$, X- finite, such that an *exchange axiom* holds:

$$\text{for all } A, B \in \mathcal{I},\ |B| < |A|, \text{ there is } x \in A - B \ \text{ s.t. } B + x \in \mathcal{I}.$$

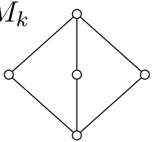
A *rank* $r(S)$ of a set $S \subseteq X$ is $r(S) = \max\{|A| \mid A \in \mathcal{I}, A \subseteq S\}$. A set S is *closed* if it is maximal of its rank.

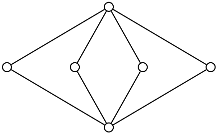
Theorem

A lattice L is geometric iff it is a lattice of closed sets of a matroid.

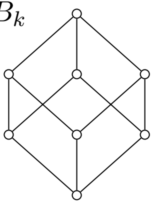
Geometric lattices (examples)

M_k

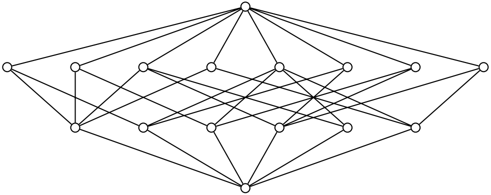




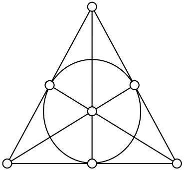
B_k



Also anything modular...



...also some weird things...



...and the Fano plane...

...and so on.

Möbius function of geometric lattices is alternating

Theorem (Rota, 1964)

For a geometric lattice L, for all $x \leq y$, $\mu(x, y)$ is nonzero and $\mathbf{sgn}\,\mu(x,y) = (-1)^{r(x)-r(y)}$.

There are many theorems about computing Möbius function. We'll need one of them:

Lemma

For 0-posets P and Q, let $p\colon P$ be an order-preserving map such that

- *for any $y \in Q$ there is $x \in P$ such that $p^{-1}[0, y] = [0, x]$;*
- *$p^{-1}(0)$ contains at least two points.*

Then for any $y \in Q$,

$$\sum_{x\colon p(x)=y} \mu(0, x) = 0.$$

Möbius function of geometric lattices is alternating

Lemma

For 0*-posets* P *and* Q*, let* $p\colon P$ *be an order-preserving mapping such that*

1. *for any* $y \in Q$ *there is* $x \in P$ *such that* $p^{-1}[0, y] = [0, x]$*;*
2. $p^{-1}(0)$ *contains at least two points.*

Then for any $y \in Q$*,*

$$\sum_{z\colon p(z)=y} \mu(0, z) = 0.$$

Proof. Note that

$$\sum_{z\colon p(z)\le y} \mu(0, z) = \left[\begin{matrix}\text{take } x = x(y) \text{ as} \\ \text{provided in (1)}\end{matrix}\right] = \sum_{z\in[0,x]} \mu(0, z)$$

$$= \left[0 < x \text{ by (1)}\right] = 0.$$

And the statement follows by induction on y. $\square$

Möbius function of geometric lattices is alternating

Lemma

For 0-posets P and Q, let $p\colon P$ be an order-preserving mapping such that

1. *for any $y \in Q$ there is $x \in P$ such that $p^{-1}[0, y] = [0, x]$;*
2. *$p^{-1}(0)$ contains at least two points.*

Then for any $y \in Q$,

$$\sum_{z\colon p(z)=y} \mu(0, z) = 0.$$

Corollary

For $a, b \in L$, $a \neq 0$, it holds:

$$\sum_{z\colon z \vee a = b} \mu(0, z) = 0.$$

Proof. Apply lemma to $P = L$, $Q = [a, 1]$ and $p = \cdot \vee a$. $\square$

Möbius function of geometric lattices is alternating

Theorem (Rota, 1964)

For a geometric lattice L, for all $x \leq y$, $\mu(x, y)$ is nonzero and $\mathbf{sgn}\,\mu(x, y) = (-1)^{r(x)-r(y)}$.

Proof. By induction: $\mu(0,0) = 1$ and $\mu(0, a) = -1$ for a - atom. Now let us take arbitrary y and $a < y$, an atom. Then (by Corollary):

$$\mu(0, y) = - \sum_{z<y:\ z \vee a = y} \mu(0, z).$$

If $z < y$ and $z \vee a = y$ then $z \not\geq a$, hence $z \wedge a = 0$. By submodularity, $r(z) + 1 = r(z) + r(a) \geq r(0) + r(y) = r(y)$. Then $r(z) = r(y) - 1$ and the statement of the theorem follows. □

Thank you!

(In case you didn't have a good look last time.)