The Möbius Function, Möbius Inversion, and their Linear Algebra

By the way, Möbius was an astronomer, although now, I believe, he's most famous as a mathematician. Funny thing, that.

In these notes we have a finite partially ordered set P and several examples, notably $\mathcal{P}(X_n)$, the power set of $X_n = \{1, 2, \dots, n\}$.

1. Matrices for Möbius

The book introduces the set $\mathcal{F}$ of all functions $\varphi : P \times P \to \mathbb{R}$ such that $\varphi(x, y) = 0$ if $x \not\leq y$. I want a name for such functions; let's call them *incidence functions* on P. That is, $\mathcal{F}$ is the set of incidence functions. Since $\mathcal{F}$ has addition and scalar multiplication of functions, and it also has multiplication, called "convolution", it is algebraic; thus we call it the *incidence algebra* of P. The book defines convolution and the functions δ, ζ, and μ.

Any incidence function φ can be thought of in terms of a matrix, which I call M_φ. The matrix is really just the table of values of φ. The rows and columns of the matrix are labelled by the elements of P, both using the same (total) ordering of elements. If that ordering is a linear extension of $<$, then *M_φ is upper triangular.* This is a very good feature for computations. Therefore, we always use a linear extension for labelling the rows and columns of the matrices.

The two properties we need of this correspondence between an incidence function and its matrix are stated in the next theorem:

Theorem 1 (Algebra of partially ordered set matrices)**.** *Let $\varphi_1, \varphi_2 \in \mathcal{F}$.*

(1) *Convolution corresponds to matrix multiplication: $M_{\varphi_1 * \varphi_2} = M_{\varphi_1} M_{\varphi_2}$.*
(2) *$M_\delta = I$, the identity matrix.*

Corollary 1.1. *If $\varphi_1 * \varphi_2 = \delta$, then $M_{\varphi_1} M_{\varphi_2} = M_\delta = I$. Therefore $M_{\varphi_1} = M_{\varphi_2}^{-1}$ and $M_{\varphi_2} = M_{\varphi_1}^{-1}$.*

In particular, $M_\mu = M_\zeta^{-1}$.

Proof. Recall that we defined μ to be the incidence function such that $\mu * \zeta = \delta$. In matrix form, therefore (by Theorem **??**), $M_\mu * M_\zeta = M_\delta = I$. Therefore $M_\mu = M_\zeta^{-1}$. (We have to know M_ζ^{-1} exists. It exists because M_ζ is upper triangular with 1's on the diagonal, therefore M_ζ is invertible. This proof that μ exists uses pure linear algebra from Math 304, without going through the computation in the book on page 186.) □

2. The Möbius function

Example 1 (The power set of 3 elements: $\mathcal{P}(X_3)$)**.** Let's compute $\mu(\varnothing, S)$ for $S \subseteq X_3$. We work up from $\mu(\varnothing, \varnothing)$, i.e., the bottom. For readability it's convenient to simplify the notation for subsets by writing, for example, $\{1,2\}$ as 12 (I will do that usually, but not always).

$$\mu(\varnothing,\varnothing) = 1,$$
$$\mu(\varnothing,3) = -\sum_{\varnothing\le x<3} \mu(\varnothing,x) = -\mu(\varnothing,\varnothing) = -1,$$

and similarly

$$\mu(\varnothing,1) = \mu(\varnothing,2) = -1.$$

Also,

$$\mu(\varnothing,23) = -\sum_{\varnothing\le x<23} \mu(\varnothing,x) = -[\mu(\varnothing,\varnothing)+\mu(\varnothing,2)+\mu(\varnothing,3)] = 1,$$

and similarly

$$\mu(\varnothing,12) = \mu(\varnothing,13) = 1.$$

Finally,

$$\mu(\varnothing,123) = -[1-1-1-1+1+1+1] = -1.$$

To compute $\mu(1,S)$ for $S\subseteq X_3$, do the same thing as above but start at $\{1\}$. Therefore you take $\mu(1,1)=1$, and find $\mu(1,12)=-1$, $\mu(1,13)=-1$, and $\mu(1,123) = -(1-1-1) = 1$. The other subsets, those with $S \not\supseteq \{1\}$, all have $\mu(\{1\},S)=0$.

The final result of calculating all values of $\mu(R,S)$ can be summarized in a general formula:

$$\mu(R,S) = \begin{cases} (-1)^{|S|-|R|} & \text{if } R\subseteq S, \\ 0 & \text{if } R\not\subseteq S, \end{cases}$$

and displayed in a table (which is the same as a matrix):

$$M_\mu = \begin{pmatrix} 1 & -1 & -1 & -1 & 1 & 1 & 1 & -1 \\ 0 & 1 & 0 & 0 & -1 & -1 & 0 & 1 \\ 0 & 0 & 1 & 0 & -1 & 0 & -1 & 1 \\ 0 & 0 & 0 & 1 & 0 & -1 & -1 & 1 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & -1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}.$$

The rows and columns of the matrix are labelled by the elements of $\mathcal{P}(X_3)$, both using the same total order of elements. I used the order $\varnothing, 1, 2, 3, 12, 13, 23, 123$. My total order is a linear extension of $<$ (which is $\subseteq$ in this example). You may have noticed that the matrix is upper triangular (not a coincidence).

There is a general theorem for the power set of an n-element set such as X_n. It is not hard to prove it by induction.

Theorem 2. *In $P(X_n)$, the Möbius function is*

$$\mu(R,S) = \begin{cases} 0 & \text{if } R \not\subseteq S, \\ (-1)^{|S|-|R|} & \text{if } R \subseteq S. \end{cases}$$

Consider a general partially ordered set P. If it has a unique maximal element we call that element $\hat{1}$. If it has a unique minimal element we call that element $\hat{0}$. That means $\hat{0} \leq x \leq \hat{1}$ for all $x \in P$.

Example 2 (A partially ordered set with $\hat{0}$ and $\hat{1}$, and 3 levels)**.** Let's write m for the number of elements in the middle level. A useful fact is that for this partially ordered set,

$$\mu(\hat{0}, \hat{1}) = -[1 + m(-1)] = m - 1.$$

The proof for small m is a homework problem.

I use this in class to deduce Steiner's 1826 counting formulas for regions and bounded regions of an arrangement of lines in the plane.

3. Möbius Inversion

Theorem 3 (Möbius Inversion, same as Theorem 6.6.1)**.** *Suppose P is a poset. Suppose $F : P \to \mathbb{R}$ is any function. Define*

$$G(x) = \sum_{y \leq x} F(y).$$

Then

$$F(x) = \sum_{y \leq x} G(y)\mu(y,x).$$

Why do we want this? Because very often we want to know F, but we can calculate G, not F; so we use Möbius inversion to get F.

Example 3. We take $P = \mathcal{P}(X_n)$. Given $F : \mathcal{P}(X_n) \to \mathbb{R}$, define

$$G(R) = \sum_{S \subseteq R} F(S).$$

Conclusion (from Theorems **??** and **??**):

$$F(R) = \sum_{S \subseteq R} G(S)\mu(S,R) = \sum_{S \subseteq R} G(S)(-1)^{|R|-|S|}.$$

Corollary 3.1. *Take $R = \varnothing$. Then $F(\varnothing) = \sum_{S \subseteq X_n} (-1)^{|S|} G(S)$.*

This corollary is what turns into the Principle of Inclusion and Exclusion. But that is off topic (it's in §6.6).

We can turn Möbius inversion into linear algebra and use linear algebra to prove Theorem **??**. Let $P = \{x_1, x_2, \dots, x_n\}$. A function $F : P \to \mathbb{R}$ can also be written as a vector:

$$F = \big(F(x_1), F(x_2), \dots, F(x_n)\big) \in \mathbb{R}^n.$$

Now we can prove Möbius inversion. We use the vector form of F in the proof.

Theorem 4 (Möbius Inversion again)**.** *Suppose we have any function $F : P \to \mathbb{R}$. If we define $G(x) = \sum_{y \le x} F(x)$, then $F(x) = \sum_{y \le x} \mu(y, x) G(y)$.*

Proof. We rewrite the definition of G with the restriction on y removed and replaced by a zeta function: $G(x) = \sum_{y \in P} F(y)\zeta(y, x)$. (Check it, to make sure you understand how this works!) We rewrite the sum as a matrix product with M_ζ on the right:

$$\begin{aligned} G &= F M_\zeta, \\ \therefore F &= G M_\zeta^{-1}, \\ \therefore F &= G M_\mu, \\ \therefore F(x) &= \sum_{y \in P} G(y)\mu(y, x), \\ \therefore F(x) &= \sum_{y \le x} G(y)\mu(y, x) \end{aligned}$$

(because μ is an incidence function so the terms where $y \not\le x$ are zero). □

Acknowledgement

Notes by Tom Zaslavsky. Thanks to Charlie DiGiovanna for offering his typed notes as a helpful first draft.