Abstract

Maple packages and data files

This article is accompanied by the maple package `RSP.txt` and several example inputs and outputs available at

`http://www.math.rutgers.edu/ãjl213/DrZ/RSP.html`

Background

There is a rich study of Dyck paths in combinatorics. Some of the most ubiquitous results are for the case that the slope of the line is 1. In particular that the number of paths from (0,0) to (n,n) is counted by the Catalan numbers. When we change it from 1 to another rational number a/b, we enter the realm of appropriately named Rational Catalan Combinatorics. For notational convenience, we'll let $A_{a,b,n}$ denote the number of paths from $(0,0)$ to (b,a) staying on or below the line $y = a/bx$.

It was shown by Duchon in 2000 that for any slope $\frac{a}{b}$, the number of paths below a line of that slope is asymptotically $\Theta\left(\frac{1}{n}\binom{(a+b)n}{an}\right)$ [D1]. However, it's still unknown what the constant out front is. To show the asymptotics, Duchon showed that the contant is somewere between $\frac{1}{a+b}$ and $\frac{1}{a}$. This upper bound on the number of paths was known at least as far back as 1950 to Grossman.

It's clear that $A_{a,b,n} = A_{b,a,n}$, so in this paper, we'll assume that we always have $a > b$.

For the case $b = 1$, there is an exact solution known, using Fuss-Catalan numbers $\frac{1}{1+an}\binom{(1+a)n}{an}$.

This Article

We would like to try and find the coefficient out front, that is, α so that the number of paths is $(1 + o(1))\frac{\alpha}{n}\binom{(a+b)n}{an}$. In an effort to do this, we first are tasked with computing many terms of the sequence

Through a simple dynamic programming algorithm, we are able to compute the number of paths from (0,0) to (bn,an) for n around a thousand. This gives us enough data to try and find a recurrence relation that it satisfies using the maple package available at `http://www.math.rutgers.edu/ zeilberg/tokhniot/FindRec.txt` However, we were only able to successfully find recurrences for the slopes 3/2 and 5/2. Armed with these recurrences we are able to blindingly fast crank out many thousands more terms of this sequence. The recurrence for slope 3/2 is given below: Though we can't guarentee this is the minimal recurrence, it still gives a massively faster of counting the paths for these two unknown slopes, and potentially for many more slopes that our computer wasn't keen enough to find this time.

Then, once we have exact numbers, we do a statistical fit of the data for many values of n against the model $\left(\frac{\alpha}{n} + \frac{\beta}{n^2} + \frac{\gamma}{n^3} + \frac{\delta}{n^4}\right)\binom{(a+b)n}{an}$ to get our estimate of α. Adding more error terms didn't affect the value of α much. We estimate how close this is to the truth by running it for the first 100 values of n, and the first 200 values of n, and seeing how much our estimate stays the same. We can get quite good estimates of these numbers!

a\b	1	2	3	4	5	6
1						
2	0.50000000	1				
3	0.333333333	0.240706636	1			
4	0.25	0.50000000	0.15972479544	1		
5	0.2	0.1613399969	0.1372518253	0.119952918	1	
6	0.166666667	0.333333333	0.50000000	0.240706636	0.09621264003	1

There are many more slopes for which we have very exact estimates of α, and are availiable online at

`http://www.math.rutgers.edu/ãjl213/DrZ/RSP.html` in the extra data folder.

A similar, but interesting and distinct problem is to try letting something else go to infinity in $A_{a,b,n}$ other than n, as we had before.

Suppose instead that we were to let a go to infinity while b is fixed. There's a nice pattern that appears. In particular, we have the conjecture that it is asymptitically equal to $\frac{gcd(a,b)}{b}$. The $gcd(a,b)$ factor makes sense because it makes the expression only depend on the fraction a/b.

Figure 1: $a\alpha$ as a function of a for $b=2$

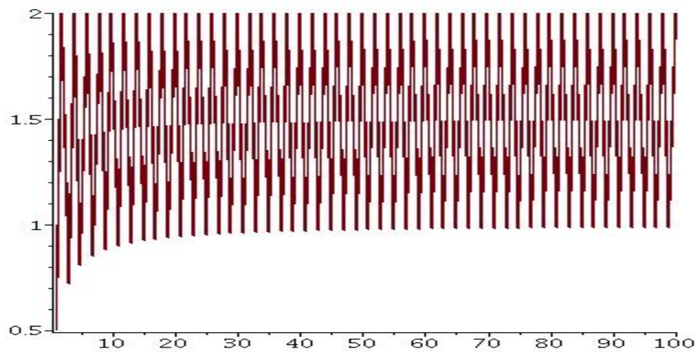
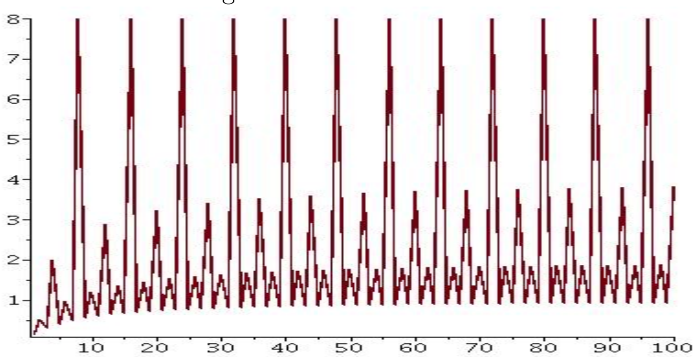


Figure 2: $a\alpha$ as a function of a for $b = 8$



A much less studied area is to consider paths in a three dimensional lattice that have to stay to stay in a region bounded by planes. It is simple to extend the dynamic programming solution to this situation. However, since there many more lattic points, the runtime goes up from $\Theta(n^2)$ to $\Theta(n^3)$. This keeps us from getting anywhere near as much data. With the data we do have, we have the suggestion that something much more interesting than in the 2D case is happening! In particular, the way that we set up the three dimensional problems, is that we take the number of paths with steps in $\{\langle 1,0,0\rangle, \langle 0,1,0\rangle, \langle 0,0,1\rangle\}$. Instead of the 2D problem of requiring $ax < by$, we define the counting problem with three numbers, a, b, c, and require of our paths that they satisfy $ax \leq by \leq cz$ that end at $x = bc, y = ac, z = ab$. If we have $a = b = c = 1$, this has the precise formula of the 3D Catalan numbers (A005789) $\frac{2}{(n+1)^2(n+2)}\binom{3n}{n,n,n}$. however, there is a lot left to understand, and some things that are distinctly different than the 2D case. In particular, for 2D, it was always $\Theta\left(\frac{1}{n}\binom{(a+b)n}{an}\right)$. That is, the slope of the line didn't affect the fact that you always had a $\Theta(\frac{1}{n})$ fraction of all paths. For the already known $a = b = c = 1$, it is a $\Theta(\frac{1}{n^3})$, however this appears to change for different choices of a, b, c. We have some data on the value of this coefficient in the following table

Table 1: a=1

b\c	1	2	3	4	5	6	7
1	3.0	2.7	2.6	2.5	2.5	2.4	2.4
2	3.7	3.3	3.0	2.9	2.8	2.7	2.6
3	4.3	3.7	3.4	3.2	3.1	3.0	2.9
4	4.8	4.1	3.8	3.5	3.4	3.2	3.1
5	5.2	4.5	4.1	3.8	3.6	3.4	3.3
6	5.7	4.8	4.4	4.1	3.8	3.6	3.5
7	6.0	5.2	4.6	4.3	4.0	3.8	3.7

Table 2: a=2

b\c	1	2	3	4	5	6	7
1	2.7	2.5	2.5	2.4	2.4	2.3	2.3
2	3.3	3.0	2.8	2.7	2.6	2.6	2.5
3	3.7	3.4	3.2	3.0	2.9	2.8	2.8
4	4.1	3.7	3.5	3.3	3.1	3.0	3.0
5	4.5	4.0	3.7	3.5	3.4	3.2	3.2
6	4.8	4.3	4.0	3.7	3.6	3.4	3.3
7	5.2	4.6	4.2	4.0	3.8	3.6	3.5

References

[D1] Phillipe Duchon, *On the enumeration and generation of generalized Dyck words*, Discrete Mathematics **225** (2000), 121-135