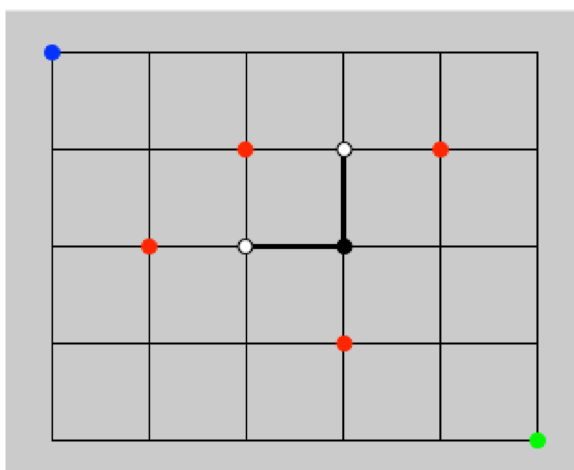
Paths in a Grid

Wagner Truppel

August 2, 2015

The problem

Suppose we have a traveller who wants to move between two points in a rectangular grid with M rows and N columns, such that he can only move to the right or downwards. Moreover, suppose that there are certain points on the grid which he wants to avoid altogether. What is the number of possible paths, given the size of the grid and the number and locations of the points to avoid? For example, in the figure below, suppose he starts at the blue dot and wants to arrive at the green dot avoiding all the red dots.



No obstructions

In order to answer this question, let's first ignore the red dots and ask how many paths there are between the blue dot (located at $(0, 0)$) and the green dot (located at (M, N)), allowing him to pass through any point in between.

To compute *that*, let $f(r,c)$ denote the number of possible paths from the starting position $(0,0)$ to the point located at (r,c) (say, the black dot), where $0 \le r \le M$ and $0 \le c \le N$. We know some simple facts about $f(r,c)$:

- For a given point at the top-most row, there is only one path from the origin to that point, so $f(0,c) = 1$ for $1 \le c \le N$.
- For a given point at the left-most column, there is only one path from the origin to that point, so $f(r,0) = 1$ for $1 \le r \le M$.
- To go from the origin to, well, the origin, we could say that there is only one path or that there is no path and both choices seem equally appropriate, so $f(0,0) = a$, where $a \in \{0,1\}$. However, to preserve continuity as either index approaches zero, it's probably best to select $a = 1$.

Now, consider the black dot in the picture. In order to arrive at that point, the traveller must have come either from the point immediately above, $(r-1,c)$, or from the point immediately to the left, $(r,c-1)$ (white dots), so the number of paths from the starting point arriving at (r,c) must equal the sum of the number of paths arriving at $(r-1,c)$ and the number of paths arriving at $(r,c-1)$. There is no overlap and no overcounting, because the traveller cannot backtrack. Thus,

$$f(r,c) = f(r-1,c) + f(r,c-1)\,, \qquad 1 \le r \le M\,, \quad 1 \le c \le N\,.$$

We are left, then, with a double recurrence relation:

$$(*) \qquad \begin{cases} f(r,c) \;=\; f(r-1,c) + f(r,c-1)\,, & 1 \le r \le M\,, \qquad 1 \le c \le N \\ f(r,0) \;=\; 1\,, & 1 \le r \le M \\ f(0,c) \;=\; 1\,, & \qquad\qquad\qquad 1 \le c \le N \\ f(0,0) \;=\; 1\,. & \end{cases}$$

By expanding the recurrence, it's not too difficult to show that

$$f(r,c) = \sum_{k=0}^{r} f(r-k,c-1) \qquad \text{and} \qquad f(r,c) = \sum_{k=0}^{c} f(r-1,c-k)\,.$$

Then, from the boundary values $f(r,0) = f(0,c) = 1$, it takes little effort to show that

$$\begin{aligned} f(r,1) &= r+1 & \qquad f(1,c) &= c+1 \\ f(r,2) &= \frac{(r+1)(r+2)}{2} & \qquad f(2,c) &= \frac{(c+1)(c+2)}{2} \\ f(r,3) &= \frac{(r+1)(r+2)(r+3)}{6} & \qquad f(3,c) &= \frac{(c+1)(c+2)(c+3)}{6} \end{aligned}$$

These suggest the result

$$f(r,c) = \frac{(r+c)!}{r!\,c!} = \binom{r+c}{r} = \binom{r+c}{c}$$

and it's not difficult to prove that it does indeed satisfy the recurrence relation and the boundary values. In retrospect, this result is quite easy to understand. In order to go from the origin to the point (r, c) by moving only forward or downwards, we have to choose — from a total of $(r + c)$ steps — r steps to move downwards and the remaining c steps to move forwards. With $f(r, c)$ at hand, we can now compute the total number of paths from $(0, 0)$ to (M, N) by simply evaluating $f(M, N)$:

$$\boxed{g(M,N) \equiv \text{total number of paths from } (0,0) \text{ to } (M,N) = \binom{M+N}{M}.}$$

Now, how can we use this result to answer the original question?

A single obstruction

To answer that, suppose first that there is a single red dot, a single obstruction, located at position (r_1, c_1). Of all the $g(M, N)$ paths that start at $(0, 0)$ and end at the destination (M, N), the number of paths that pass through the obstruction is the number of paths that start at $(0, 0)$ and end at the obstruction plus the number of paths that start at the obstruction and end at the destination (M, N).

The number of paths that start at $(0, 0)$ and end at the obstruction at (r_1, c_1) can be computed as if we had an unobstructed grid of size (r_1, c_1), so it's $g(r_1, c_1)$. Similarly, the number of paths that start at the obstruction and end at the destination (M, N) can be computed as if we had an unobstructed grid of size $(M - r_1, N - c_1)$, so it's $g(M - r_1, N - c_1)$. Thus, the number of possible paths from $(0, 0)$ to (M, N) that avoid the single obstruction at (r_1, c_1) is given by:

$$\boxed{h_{M,N}\big[\,(r_1,c_1)\,\big] = g(M,N) - \Big(g(r_1,c_1) + g(M-r_1,N-c_1)\Big)}$$

Two obstructions

What if we have two obstructions (red dots), say, at (r_1, c_1) and at (r_2, c_2)? Things now get more complicated. If there are no possible paths connecting the two obstructions, then the number of possible paths from $(0,0)$ to (M,N) that avoid both obstructions is simply

$$\begin{aligned} h_{M,N}\big[\,(r_1,c_1),(r_2,c_2)\,\big] \;&=\; g(M,N) \\ &-\; \Big(g(r_1,c_1)+g(M-r_1,N-c_1)\Big) \\ &-\; \Big(g(r_2,c_2)+g(M-r_2,N-c_2)\Big) \end{aligned}$$

but if there are paths connecting the two obstructions then the above subtracted too many paths, namely, those that start at $(0,0)$, pass through both obstructions, then end at (M,N), and we need to add them back once. Thus, if the obstructions are connectable by valid paths, then

$$\begin{aligned} h_{M,N}\big[\,(r_1,c_1),(r_2,c_2)\,\big] \;&=\; g(M,N) \\ &-\; \Big(g(r_1,c_1)+g(M-r_1,N-c_1)\Big) \\ &-\; \Big(g(r_2,c_2)+g(M-r_2,N-c_2)\Big) \\ &+\; \Big(g(r_1,c_1)+g(r_2-r_1,c_2-c_1)+g(M-r_2,N-c_2)\Big) \end{aligned}$$

or, after a trivial simplification,

$$h_{M,N}\big[\,(r_1,c_1),(r_2,c_2)\,\big] = g(M,N)-g(M-r_1,N-c_1)-g(r_2,c_2)+g(r_2-r_1,c_2-c_1)$$

where we assumed that the two obstructions are sorted such that $r_2 \geq r_1$ and $c_2 \geq c_1$, which is always possible if they are connectable by valid paths. In fact, that *is* the condition that the two obstructions must satisfy for them to be connectable because when $r_2 \geq r_1$ and $c_2 < c_1$ (or $c_2 \geq c_1$ and $r_2 < r_1$) the traveller cannot backtrack from one obstruction to the other. However, if $r_2 \geq r_1$ and $c_2 \geq c_1$, the second obstruction is always ahead of the first obstruction and can be reached from it.

Thus, in summary, when there are two obstructions, at (r_1,c_1) and at (r_2,c_2), the number of paths from $(0,0)$ to (M,N) which avoid both obstructions is given by

- if $r_2 \geq r_1$ and $c_2 \geq c_1$ (second obstruction reachable from the first):

$$\boxed{h_{M,N}\big[\,(r_1,c_1),(r_2,c_2)\,\big] = g(M,N)-g(M-r_1,N-c_1)-g(r_2,c_2)+g(r_2-r_1,c_2-c_1)}$$

- if neither obstruction is reachable from the other:

$$\boxed{\begin{aligned} h_{M,N}\big[\,(r_1,c_1),(r_2,c_2)\,\big] \;&=\; g(M,N) \\ &-\; \Big(g(r_1,c_1)+g(M-r_1,N-c_1)\Big) \\ &-\; \Big(g(r_2,c_2)+g(M-r_2,N-c_2)\Big) \end{aligned}}$$

Three or more obstructions

...

∎