Counting Kings: Explicit Formulas, Recurrence Relations, and Generating Functions! Oh My!

Neil J. Calkin, Kevin James, Shannon Purvis, Shaina Race, Clemson; Keith Schneider, UNCA; Matthew Yancey, Virginia Tech

July 18, 2006

Abstract

Let F(m, n) be the number of distinct configurations any number of non-attacking kings on an $m \times n$ chessboard. We'll explore the sequences generated when we fix n at some small value and discuss methods of finding explicit formulas, recurrence relations and generating functions for these sequences.

1 Introduction

Consider the function in two variables $F(m,n) = \mathbf{1}^T A_n^{m-1} \mathbf{1}$ where A_n is a recursively defined matrix as follows:

$$A_0 = (1), A_1 = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}, \text{ and } A_n = \begin{pmatrix} A_{n-1} & A_{n-2} \\ A_{n-2} & 0 \end{pmatrix}$$

with the copies of A_{n-2} padded with zeros as needed and 0 being a zero matrix of the appropriate size. Additionally, $\mathbf{1}^T$ and $\mathbf{1}$ are respectively a row and column vector of all ones of the necessary length for multiplicative conformity. In other words F(m, n) is the sum of the elements of A_n raised to the m-1 power.

F(m, n) gives the number of distinct configurations of an arbitrary number of non-attacking kings on an $m \times n$ chessboard. This was derived using techniques found in Calkin and Wilf[2] but proof of this will not be presented herein. F(m, n) also counts the distinct tilings of an $(m+1) \times (n+1)$ rectangle with 1×1 and 2×2 tiles which is a problem studied by Heubach[3].

This research was funded, in part, by a grant from the NSF.

A simple bijection between these two problems can be found by placing a king in the same corner of every 2×2 tile and removing the appropriate edges. While the derivation of this formula will not be discussed it is useful to note that A_n is an adjacency matrix of a graph with each vertex corresponding to a legal configuration of kings on an $1 \times n$ chessboard and edges between vertices signifying which of these $1 \times n$ boards can be placed adjacent to each other in the construction of an $m \times n$ board.

A little experimentation reveals that F(m, 1) yields the familiar Fibonacci sequence properly indexed, and F(m, 2) gives the less familiar Jacobsthal sequence. Proof of each of these is a simple exercise as is finding explicit formulas, recurrence relations, and generating functions. What follows are methods for finding explicit formulas, recurrence relations, and generating functions for other small fixed n.

2 Explicit formulas for F(m, n) with fixed n

There are already well known explicit formulas for F(m, 1) and F(m, 2). In this section we will discuss the derivation of explicit formulas when we fix n at some other value.

Since each A_n is a real, symmetric matrix the principle of spectral decomposition as found in Axler[1] allows us to state:

$$A_n^{m-1} = \sum \lambda_{n,i}^{m-1} \frac{e_{n,i} e_{n,i}^T}{||e_{n,i}||^2}$$

Where each $\lambda_{n,i}$ is an eigenvalue of A_n with associated eigenvector $e_{n,i}$. Note that $e_{n,i}$ is a column eigenvector while $e_{n,i}^T$ is a row eigenvector so $e_{n,i}e_{n,i}^T$ is a square matrix with the same dimension as A_n .

We are primarily concerned with the sum of the elements of A_n^{m-1} and have previously stated the formula:

$$F(m,n) = \mathbf{1}^T A_n^{m-1} \mathbf{1}$$

So we can get the same result by summing the elements of each matrix in our decomposition:

$$F(m,n) = \sum \lambda_{n,i}^{m-1} \frac{\mathbf{1}^T e_{n,i} e_{n,i}^T \mathbf{1}}{||e_{n,i}||^2}$$

Notice that $\mathbf{1}^T e_{n,i}$ and $e_{n,i}^T \mathbf{1}$ are each the sum of the elements of the $e_{n,i}$ so $\mathbf{1}^T e_{n,i} e_{n,i}^T \mathbf{1}$ is the square of the sum of the elements of $e_{n,1}$. Additionally $||e_{n,i}||^2$ is the sum of the square of those elements. Thus, $\frac{\mathbf{1}^T e_{n,i} e_{n,i}^T \mathbf{1}}{||e_{n,i}||^2}$ is a

real, non-negative number since all the elements of every eigenvector of a real, symmetric matrix are real numbers.

We can see that zero eigenvalues contribute nothing to this sum since each summand has $\lambda_{n,i}$ as a factor. However, each summand also has a factor of $\mathbf{1}^T e_{n,i} e_{n,i}^T \mathbf{1}$ which is the square of the sum of the elements of the eigenvalue. This means that eigenvalues with associated eigenvectors which have entries that sum to zero also do not contribute to our sum. The following two lemmas will give us a way to find the eigenvalues which do contribute to the sum.

Lemma 1. The first entry of an eigenvector with a zero eigenvalue is zero.

Proof. Let v be an eigenvector of A_n associated with a zero eigenvalue with initial entry v_1 . By our method of construction of A_n we know it has at least one row that begins with 1 and has all other entries 0. We'll call this the *kth* row and note that the dot product of the *kth* row and v is v_1 corresponding to the *kth* entry of $A_n v$. However, since v is associated with a zero eigenvalue, all entries of $A_n v$ must be zero. Therefore, $v_1 = 0$.

Lemma 2. With v an eigenvector of A_n , the initial element of v is equal to zero iff the sum of the elements of v is zero.

Proof. \Rightarrow Let v be an eigenvector with initial entry zero associated with an eigenvalue, λ . Since the first row of A_n is a row of all ones, the first element of $A_n v$ is the sum of the elements of v. Additionally, the first entry of $A_n v$ is $\lambda 0 = 0$. Therefore the sum of the elements of v must be zero.

 \Leftarrow Now let v be an eigenvector of A_n with elements that sum to zero associated with an eigenvalue, λ . Let v_1 be the first element of v. So the first element $A_n v$ is λv_1 , but this is also the sum of the elements of v as above which is zero. So if $\lambda \neq 0$, v_1 must be zero and if $\lambda = 0$ we have $v_1 = 0$ from the previous lemma.

Using the previous two lemmas we can see that $\lambda_{n,i}^{m-1} \frac{\mathbf{1}^T e_{n,i} e_{n,i}^T}{||e_{n,i}||^2} \neq 0$ precisely when the initial entry of $e_{n,i} \neq 0$. We will call the subset of the eigenvalues of A_n for which this is true the active eigenvalues. The non-zero eigenvalues that are not active will be called inactive eigenvalues.

With this in mind let:

$$v = \begin{pmatrix} 1 \\ x_2 \\ \vdots \\ x_{d-1} \\ x_d \end{pmatrix}$$

Where d is the dimension of A_n . Using the system given by $A_n v = \lambda v$ we can solve for x_2, x_3, \ldots, x_d each in terms of λ . Since the sum of the elements

of v must equal λ we have the equation $1 + x_2 + x_3 + \ldots + x_d = \lambda$. We can then transform this equation into a polynomial, P_n , the roots of which are exactly the active eigenvalues of A_n . We can also define $R_n(\lambda) = \frac{1^T v v^T 1}{||v||^2}$ and $k_{n,i} = R_n(\lambda_{n,i})$. Now we have $F(m,n) = \sum k_{n,i} \lambda_{n,i}^{m-1}$ for each $\lambda_{n,i}$ that is a root of P_n . Also, since there are a finite number of active eigenvalues of A_n for each value of n we have an explicit formula for F(m,n) with finitely many terms when we fix n.

For clarity, we'll calculate the explicit formula for F(m,3). Using our definition of A_n we find:

Let:

$$v = \begin{pmatrix} 1\\ x_2\\ x_3\\ x_4\\ x_5 \end{pmatrix}$$

The matrix equation $A_3v = \lambda v$ gives the following system of equations:

$$\begin{array}{rclrcl}
1 + x_2 + x_3 + x_4 + x_5 &=& \lambda \\
& 1 + x_4 &=& \lambda x_2 \\
& 1 &=& \lambda x_3 \\
& 1 + x_2 &=& \lambda x_4 \\
& 1 &=& \lambda x_5
\end{array}$$

Solving the last four equations for x_2, x_3, x_4, x_5 in terms of λ yields:

$$x_2 = \frac{1}{\lambda - 1}, \ x_3 = \frac{1}{\lambda}, \ x_4 = \frac{1}{\lambda - 1}, \ x_5 = \frac{1}{\lambda}$$

And substituting these into the first equation gives:

$$\begin{split} 1 + \frac{1}{\lambda - 1} + \frac{1}{\lambda} + \frac{1}{\lambda - 1} + \frac{1}{\lambda} &= \lambda \\ \Rightarrow \lambda(\lambda - 1) \left(1 + \frac{1}{\lambda - 1} + \frac{1}{\lambda} + \frac{1}{\lambda - 1} + \frac{1}{\lambda} \right) &= \lambda(\lambda - 1)\lambda \\ \Rightarrow \lambda^2 + 3\lambda - 2 &= \lambda^3 - \lambda^2 \\ \Rightarrow \lambda^3 - 2\lambda^2 - 3\lambda + 2 &= 0 \end{split}$$

This means $P_3 = \lambda^3 - 2\lambda^2 - 3\lambda + 2$ and the three real roots of P_3 are our active eigenvalues. We also have:

$$v = \begin{pmatrix} 1\\ \frac{1}{\lambda - 1}\\ \frac{1}{\lambda}\\ \frac{1}{\lambda - 1}\\ \frac{1}{\lambda} \end{pmatrix}$$

We can use this to calculate $R_3 = \frac{\mathbf{1}^T v v^T \mathbf{1}}{||v||^2}$ in terms of λ and we'll also use the fact that the numerator of this fraction is the square of the sum of the elements of v and that sum is λ . So:

$$R_3 = \frac{\lambda^2}{1 + \frac{1}{(\lambda - 1)^2} + \frac{1}{\lambda^2} + \frac{1}{(\lambda - 1)^2} + \frac{1}{\lambda^2}} = \frac{\lambda^4 (\lambda - 1)^2}{\lambda^4 - 2\lambda^3 + 5\lambda^2 - 4\lambda + 2}$$

Finally we have:

$$F(m,3) = \sum_{i=1}^{3} k_{3,i} \lambda_{3,i}^{m-1} = k_{3,1} \lambda_{3,1}^{m-1} + k_{3,2} \lambda_{3,2}^{m-1} + k_{3,3} \lambda_{3,3}^{m-1}$$

Where each $\lambda_{3,i}$ is a root of P_3 and each $k_{3,i} = R_3(\lambda_{3,i})$. Unfortunately, the roots of P_3 are not rational so we are limited to a numerical approximation where $\lambda_{3,1} \approx 2.81361$, $\lambda_{3,2} \approx -1.34292$, $\lambda_{3,3} \approx 0.529317$, $k_{3,1} \approx 4.25452$, $k_{3,2} \approx 0.729154$, and $k_{3,3} \approx 0.0163216$.

It is interesting to note that the dimension of A_3 is 5 and the rank of A_3 is 4, yet there are only 3 active eigenvalues. $(A_1 \text{ and } A_2 \text{ both have rank } 2 \text{ and } 2 \text{ active eigenvalues.})$ A_3 has -1 as an eigenvalue with corresponding eigenvector $(0 \ 1 \ 0 \ -1 \ 0)$, and it is easy to see that the sum of the elements of that eigenvector is, in fact, zero. This verifies the existence of inactive eigenvalues, and it appears that every A_n with $n \ge 3$ has inactive eigenvalues. If we define r_n^* to be the number of active eigenvalues of A_n we can state that $r_n^* \le r_n$ where r_n is the rank of A_n . Additionally, by our method of construction of P_n we know that P_n is a factor of the characteristic polynomial of A_n and the degree of P_n is equal to r_n^* . The sequence given by r_n^* for $1 \le n \le 16$ is:

 $2,\ 2,\ 3,\ 4,\ 6,\ 8,\ 14,\ 19,\ 32,\ 45,\ 70,\ 103,\ 162,\ 241,\ 376,\ 565$

The next term in this sequence is not known. In contrast the sequence given by r_n for the same values of n is:

2, 2, 4, 6, 10, 14, 22, 34, 54, 82, 126, 194, 302, 466, 718, 1106

So we can see that there are quite a few inactive eigenvalues in general.

We have included a table of P_n and R_n for small values of n. We have calculated P_n for $n \leq 12$ and R_n for $n \leq 9$ but will not include all of them here as the equations are lengthy.

n	P_n
1	$\lambda^2 - \lambda - 1$
2	$\lambda^2 - \lambda - 2$
3	$\lambda^3 - 2\lambda^2 - 3\lambda + 2$
4	$\lambda^4 - 2\lambda^3 - 7\lambda^2 + 2\lambda + 3$
5	$\lambda^6 - 2\lambda^5 - 16\lambda^4 - \lambda^3 + 27\lambda^2 - \lambda - 4$

Table 1: P_n for $n \leq 5$

n	R_n
1	$rac{\lambda^4}{\lambda^2+1}$
2	$rac{\lambda^4}{\lambda^2+2}$
3	$\frac{\lambda^4(\lambda-1)^2}{\lambda^4-2\lambda^3+5\lambda^2-4\lambda+2}$
4	$rac{\lambda^4(\lambda^2-\lambda-1)^2}{\lambda^6-2\lambda^5+6\lambda^4+6\lambda+3}$
5	$\frac{\lambda^4(\lambda^4-\lambda^3-5\lambda^2+1)^2}{\lambda^{10}-2\lambda^9+3\lambda^8+22\lambda^7+3\lambda^6+50\lambda^5+104\lambda^4-28\lambda^3-33\lambda^2+4}$

Table 2: R_n for $n \leq 5$

3 Recurrence relations for F(m, n)

Since $F(m,n) = \sum_{i=1}^{r_n^*} k_{n,i} \lambda_{n,i}^{m-1}$ for any fixed *n* there must exist a recurrence relation for F(m,n) consisting of r_n^* terms. If we have P_n , as defined in the previous section, in hand we can compute the coefficients of the recurrence relation easily owing to the fact that P_n is the characteristic equation of the desired recurrence relation.

For example, P_1 is $\lambda^2 - \lambda - 1$. Setting $P_n = 0$ and solving for the highest degree term yields $\lambda^2 = \lambda + 1 \Rightarrow F(m, 1) = F(m - 1, 1) + F(m - 2, 1)$. We can perform a similar operation on any P_n . However, P_n is difficult to compute, and we may not have it at our disposal. We also may not know r_n^* as based on the material discussed so far we must essentially compute the entire eigensystem of A_n in order to find r_n^* .

First we will outline an easier method for finding r_n^* for any fixed n. Using $F(m,n) = \mathbf{1}^T A_n^{m-1} \mathbf{1}$, generate the first $2r_n - 1$ terms of F(m,n) where r_n is the rank of A_n . Populate a $r_n \times r_n$ matrix with these values using F(1,n) to F(r,n) in order as the first row, F(2,n) to F(r+1,n) as the second and so on. The rank of this matrix is r_n^* . This is guaranteed to work since $r_n^* \leq r_n$. In actual practice r_n^* is much less than r_n so we can get away with using a smaller matrix. However, this has not been proved to be true. We can, however, guess a number R that should be slightly bigger than r_n^* and generate matrices as above of dimension R and R+1. If these two matrices have the same rank, that value is r_n^* , if they do not, their ranks are R and R+1 which means our guess was too low.

Now we'll create a similar matrix, B, that is $r_n^* \times r_n^*$ placing the values

F(1,n) to $F(2r_n^*-1,n)$ as above. Let v be a column vector with r_n^* elements $\alpha_1, \ldots, \alpha_{r_n^*}$ with indices increasing from bottom to top, and let c be a column vector with elements $F(r_n^*+1,n)$ to $F(2r_n^*,n)$ entered from top to bottom. Then solve the system Bv = c for $\alpha_1, \ldots, \alpha_{r_n^*}$. The recurrence relation is:

$$F(m,n) = \sum_{i=1}^{r_n} \alpha_i F(m-i,n)$$

Unfortunately, there is a major drawback to this approach. For all but the smallest values of n, F(m, n) grows so rapidly that solving the resulting system takes a considerable amount of time. For example, if we wish to compute the recurrence relation for F(m, 8) we need to calculate up to F(38, 8) which is an integer with 42 digits. Solving a system with numbers this large is difficult, and it only gets worse from there. F(64, 9) has 79 digits.

There is a way to ameliorate this problem, but not to rid ourselves of it entirely. Let X_i be the number of legal configurations of kings on an $(i-1) \times n$ board that contains at least one king in each row. Then there are $X_iF(m-i,n)$ legal configurations where the first row with no kings in it is the *ith* row. Furthermore let Y be the number of configurations of kings on an $m \times n$ board that have at least one king in every row. Now we have $F(m,n) = (\sum_{i=1}^{m} X_i F(m-i,n)) + Y$. Additionally, it is obvious that $X_1 = 1$ and X_2 is one less than the dimension of A_n or alternatively $X_2 = F(1,n) - 1$. We can use a submatrix of A_n which we'll call B_n to calculate X_i for $i \ge 2$. B_n is formed by removing the first row and column of A_n and then, if we wish, we can remove rows and columns that only contain zeros. This last step is not necessary but does yield a smaller matrix. Effectively we are removing the vertices from our graph that correspond to the row with no kings and rows that can only be adjacent to the row with no kings.

For example:

Now we'll define a function; let G(m, n) be the number of configurations of non-attacking kings on an $m \times n$ that have a king in every row. For $m \geq 2$ and $n \geq 3$, $G(m,n) = \mathbf{1}^T B_n^{m-1} \mathbf{1}$ using the same logic that allowed our matrix definition of F(m,n). So $G(i-1,n) = X_i$ for $i \geq 2$ and G(m,n) = Y. This yields:

$$F(m,n) = F(m-1,n) + (F(1,n)-1)F(m-2,n)$$
$$+ \left(\sum_{i=2}^{m-1} G(i,n)F(m-i-1,n)\right) + G(m,n)$$

There are certainly similarities between the behavior of B_n and A_n . The rank of B_n must be less than the rank of A_n since the row we removed can not be involved in any linear dependencies. B_n also has active and inactive eigenvalues. In fact the inactive eigenvalues of A_n are all also inactive eigenvalues of B_n . We know this because the initial entry of every inactive eigenvector of A_n is zero. Removing the first column of all ones and removing the initial entry of the vector have no effect on the interaction of an inactive eigenvector and the rest of the matrix. These facts taken together mean that the number of active eigenvalues of B_n is less than r_n^* .

With this in mind we can compute the recurrence relation for G(m, n) using the method described above. This is easier than doing the same computation for F(m, n) for two reasons, first the recurrence has less terms, but this is a minor advantage, and secondly, and much more importantly, G(m, n) grows much more slowly than F(m, n) which leads to computations with numbers that are not quite so enormous.

Now assume we have the recurrence relation for G(m, n) and it has k terms. Thus:

$$G(m,n) = \sum_{i=1}^{k} \beta_i G(m-i,n)$$

We will now rewrite our new formula for F(m, n). However, the formula is very long so we will make use of some abbreviations. Since it is assumed that we have fixed a values of n, we can abbreviate F(m, n) to F_m and G(m, n) to G_m without really losing any information. Then:

$$F_m = F_{m-1} + (F_1 - 1)F_{m-2} + \left(\sum_{i=2}^{m-1} G_i F_{m-i-1}\right) + G_m$$

We can, of course, shift indices in order to write formulas for F_{m-1}, F_{m-2}, \ldots . Using these formulas we can write an expression for:

$$F_m - \sum_{i=1}^k \beta_i F_{m-i}$$

Solving this expression for F_m gives us:

$$F_m = \left(\sum_{i=1}^m \alpha_i F_{m-i}\right) + \alpha_{m+1}$$

Which is a recurrence relation with more terms than we desired. However, consider α_{k+3} . Our expression for F_m had the coefficient of the F_{m-k-3} term as G_{k+2} . Our expression for $-\beta_1 F_{m-1}$ has the same coefficient of the F_{m-k-3} term as $-\beta_1 G_{k+1}$. If we continue along these lines we see that:

$$\alpha_{k+3} = G_{k+2} - \sum_{i=1}^{k} \beta_i G_{k+2-1} = 0$$

This is also true for all α_j where j > k+3. Thus $r_n^* = k+2$ and we can compute every α if we know only the coefficients of the first k+2 terms of:

$$F_m = F_{m-1} + (F_1 - 1)F_{m-2} + \left(\sum_{i=2}^{m-1} G_i F_{m-i-1}\right) + G_m$$

We'll compute the recurrence relation for F(m, 4) here to demonstrate this. Recall:

$$B_4 = \left(\begin{array}{rrrrr} 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \end{array}\right)$$

Using this matrix to find the first few terms of G(m, 4) we find that G(m, 4) = G(m-1, 4) + G(m-2, 4) and k = 2. This means we only need the first four terms of our expression for F(m, 4) but we will include the fifth to show that it does, in fact, zero out. First we compute G(2, 4) = 6, G(3, 4) = 10, and G(4, 4) = 16. Then:

$$F_m = F_{m-1} + 7_{m-2} + 6F_{m-3} + 10F_{m-4} + 16F_{m-5} + \dots$$

$$-F_{m-1} = -F_{m-2} - 7F_{m-3} - 6F_{m-4} - 10F_{m-5} + \dots$$

$$-F_{m-2} = -F_{m-3} - 7F_{m-4} - 6F_{m-5} + \dots$$

Adding these three equations together gives:

$$F_m - F_{m-1} - F_{m-2} = F_{m-1} + 6F_{m-2} - 2F_{m-3} - 3F_{m-4} + 0$$
$$\Rightarrow F_m = 2F_{m-1} + 7F_{m-2} - 2F_{m-3} - 3F_{m-4}$$

Which we know to be correct by examining P_4 .

Included below is a table containing the coefficients of the first few recurrence relations. These have been computed up to n = 12 but are not all included here.

n	$\alpha_1, \ldots, \alpha_{r_n^*}$
1	1, 1
2	1, 2
3	2, 3, -2
4	2, 7, -2, -3
5	2, 16, 1, -27, 1, 4
6	3, 30, -17, -138, 85, 116, -42, -32

Table 3: α_i where $F(m, n) = \sum \alpha_i F(m - i, n)$

4 Generating functions for F(m, n)

With a recurrence relation and the first few terms of a sequence we can use the method outlined by Wilf[5] to find a generating function. The first few generating functions are given in a table below. Notice the similarity between the denominator of the *nth* generating function and P_n (i.e. the denominator is $x^{r_n^*}P_n(1/x)$).

n	g.f.(F(m,n))
1	$\frac{1+x}{1-x-x^2}$
2	$\frac{1+2x}{1-x-2x^2}$
3	$\frac{1+3x-2x^2}{1-2x-3x^2+2x^3}$
4	$\frac{1+6x-2x^2-3x^3}{1-2x-7x^2+2x^3+3x^4}$
5	$\frac{1+11x+x^2-26x^3+x^4+4x^5}{1-2x-16x^2-x^3+27x^4-x^5-4x^6}$

Table 4: g.f.(F(m,n)) for $n \leq 5$

5 Conclusion

We had first hoped that calculating the first few explicit formulas, recurrence relations, or generating functions would reveal a pattern that would allow us to predict the subsequent formulas. Unfortunately, if such a pattern exists it is elusive to say the least. However, the sequences generated when we fix n do appear to be of general interest as many of them were already included in Sloane[4] in relation to the tiling problem mentioned in the introduction, and while many of the entries in Sloane listed the first few terms, we were able to add recurrence relations to most of these entries.

6 Acknowledgments

The authors would like to thank the NSF for providing the funding that brought our group together at the REU at Clemson. We'd also like to thank Professor Mark McClure of UNCA for answering fractal related questions and helping with Mathematica[®]. Additionally, Jonathan Johnson of Clemson provided invaluable assistance on the numerous and varied programming problems we faced. Finally we'd like to thank Tim Flowers of Clemson for help with $L^{AT}EX$.

References

- [1] Sheldon Axler, Linear Algebra Done Right. Springer Verlag, 1997.
- [2] N.J. Calkin and H.S. Wilf, The number of independent sets in a grid graph, SIAM J. Discrete Math, 11 (1998) 54-60.
- [3] S. Heubach, Tiling an n-by-m Area with Squares of Size up to k-by-k (m ≤ 5), Congressus Numerantium, 140 (1999), pp. 43-64.
- [4] N. J. A. Sloane, The On-Line Encyclopedia of Integer Sequences, http://www.research.att.com/~njas/sequences/.
- [5] Herbert S. Wilf, generatingfunctionology. 3rd edition, A K Peters, 2005.