A SHORT PROOF OF THE RANK FORMULA FOR INCLUSION MATRICES

LIAM JOLLIFFE

ABSTRACT. We present a new proof of the well known formula for the rank of the inclusion matrix by constructing a $k\mathfrak{S}_m$-module spanned by the columns of this matrix and calculating its dimension. This gives a representation theoretic interpretation of this formula as the sum of dimensions of Specht modules.

1. Introduction

The *inclusion matrix*, $A_i^n(m)$, where $i \leq n \leq m$, is the $\binom{m}{i} \times \binom{m}{n}$ matrix whose rows are indexed by subsets of $[m] := \{1, 2, \ldots, m\}$ of size i and whose columns are indexed by subsets of $[m]$ of size n. The entry corresponding to position X, Y is 1 if $X \subseteq Y$ and 0 otherwise. This matrix arises in a number of combinatorial investigations. Gottlieb proved that over a field of characteristic 0 this matrix has full rank [4]. Linial and Rothschild then determined a formula for the rank of this matrix over the field of two elements, as well the special case when $n = i + 1$ over the field of three elements [7]. Wilson solved the problem over any field by proving the following [8]:

Theorem 1. *Let k be a field of characteristic p and suppose $i \leq \min\{n, m - n\}$. Then*

$$rank_k(A_i^n(m)) = \sum_{p \nmid \binom{n-j}{i-j}} \binom{m}{j} - \binom{m}{j-1},$$

where $\binom{m}{-1}$ is interpreted as 0.

Wilson also gives a characterisation of those vectors which are in the $\mathbb{Z}$-span of the columns of $A_i^n(m)$. There are other proofs of Theorem 1 in the litereature, Frankl [2], [3]. The paper of Frumking and Yakir [3] also uses representation theoretic techniques, however their proof relies on constructing idempotent elements and does not work over fields of characteristic two. Our representation theoretic proof is simpler and works in all characteristics, including fields of characteristic 0, which gives a representation theoretic interpretation of the orriginal work of Gottlieb. Observe that there is nothing lost by the assumption that $i \leq \min\{n, m - n\}$, because $A_i^n(m)^T = A_{m-n}^{m-i}(m)$, and so this assumption shall be made throughout.

We shall give a new proof of Theorem 1 by constructing a $k\mathfrak{S}_m$-module spanned by the columns of $A_i^n(m)$ which, of course, has dimension $rank_k(A_i^n(m))$. This proof shall make use of the representation theory of the symmetric group, which we review in the next section. The reader is referred to James' book [5], from which our notation is taken, for more details.

2. Representation Theory of $\mathfrak{S}_m$

Recall, a *partition* of an integer, $m > 0$, is a non-increasing sequence of positive integers $\lambda = (\lambda_1, \lambda_2, \ldots, \lambda_r)$ with $\sum_{i \geq 1} \lambda_i = m$. If λ is a partition of n we write $\lambda \vdash m$ and identify λ with its corresponding *Young diagram*: a left justified array of boxes with λ_i boxes in the ith row. Partitions of m index a number of important classes of $k\mathfrak{S}_m$-modules; we shall now describe two such classes of modules: the permutation modules M^λ and the Specht modules S^λ. To do so we will need some more combinatorial definitions.

Given a partition λ, we define a λ-*tableau* to be a bijection between the set $[m]$ and the boxes of (the Young diagram of) λ. We can define an equivalence relation $\sim_r$ on the set of λ-tableaux by calling t and s *row-equivalent* if the set of elements which appear in each row of t and s are the same. For example; the following two $(4,3,3,1)$-tableaux are row equivalent.

$$
\begin{array}{|c|c|c|c|}
\hline
1 & 2 & 3 & 4 \\
\hline
5 & 6 & 7 \\
\cline{1-3}
8 & 9 & 10 \\
\cline{1-3}
11 \\
\cline{1-1}
\end{array}
\quad \sim_r \quad
\begin{array}{|c|c|c|c|}
\hline
2 & 4 & 3 & 1 \\
\hline
6 & 5 & 7 \\
\cline{1-3}
9 & 10 & 8 \\
\cline{1-3}
11 \\
\cline{1-1}
\end{array}
$$
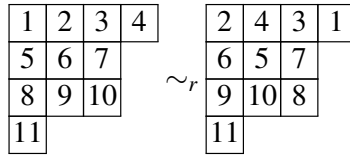
We call an equivalence class of λ-tableaux a λ-*tabloid*, and denote the tabloid corresponding to the tableaux t by $\{t\}$, or by drawing the Young diagram without the vertical lines,

$$
\begin{array}{cccc}
\hline
1 & 2 & 3 & 4 \\
\hline
5 & 6 & 7 \\
\hline
8 & 9 & 10 \\
\hline
11 \\
\hline
\end{array}
\;.
$$
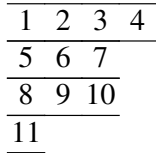
There is an obvious action of $\mathfrak{S}_m$ on both the sets of λ-tableaux and λ-tabloids, obtained by permuting the positions in which elements appear. Thus, the vector space consisting of formal sums of λ-tabloids is a $k\mathfrak{S}_m$-module, which we denote M^λ. Throughout this paper we shall only be interested in such a module when λ is a *two-part partition*, that is $\lambda = (\lambda_1, \lambda_2)$, in which case we can identify the tabloid $\{t\}$ with the set of elements appearing in the second row of $\{t\}$. We then see that $M^{(m-i,i)}$ has a basis consisting of all the i subsets $X \subseteq_i [m]$, that is subsets of size i. We shall frequently alternate between these two notations depending on notational convenience.

An important submodule of this permutation module M^λ is the Specht module S^λ, which is spanned by the polytabloids in M^λ. We shall take a slightly unusual step here, and define the Specht module as a special case in a larger family of submodules. We will specialise here to two part partitions, but the general definitions can be found in James' book [5].

Let t be a $(m-i,i)$-tableau and let $0 \leq j \leq i$. The *j-column stabiliser* of t, denoted $C_j(t)$ is the set of permutations that fix all but the first j columns of t and only permute elements that appear in the same column of t. In particular $C_j(t)$ is generated by the j transpositions which swap an element of the first j entries in the first row of t with the element appearing below it.

The *j-column symmetriser* is the element of the group algebra

$$
\kappa_j(t) := \sum_{\sigma \in C_j(t)} (-1)^\sigma \sigma,
$$

and the *j-polytabloid* is

$$e_t^j = \kappa_j(t)\{t\} \in M^{(m-i,i)},$$

where $(-1)^\sigma$ is the sign of the permutation σ. Note in particular that

$$e_t^0 = \{t\}.$$

We define the $k\mathfrak{S}_m$-module $S^{(m-i,j)(m-i,i)} \subseteq M^{(m-i,i)}$ to be the submodule spanned by the j-polytabloids. The Specht-module $S^{(m-i,i)}$ is just the module obtained when $j = i$. This gives us a chain of submodules

$$M^{(m-i,i)} \supseteq S^{(m-i,1)(m-i,i)} \supseteq \cdots S^{(m-i,i-1)(m-i,i)} \supseteq S^{(m-i,i)} \supseteq 0.$$

It can be shown that the successive quotients $S^{(m-i,j)(m-i,i)}/S^{(m-i,j+1)(m-i,i)}$ are isomorphic as $k\mathfrak{S}_m$-modules to the Specht module $S^{(m-j,j)}$, and so we write

$$M^{(m-i,i)} \sim \begin{array}{c} S^{(m)} \\ S^{(m-1,1)} \\ \vdots \\ S^{(m-i,i)} \end{array}$$

to indicate that $M^{(m-i,i)}$ has a chain of submodules whose successive quotients are $S^{(m)}, S^{(m-1,1)}, \ldots, S^{(m-i+1,i-1)}$ and $S^{(m-i,i)}$. To see that these quotients are indeed the Specht modules we follow [6] and define a $k\mathfrak{S}_m$-homomorphism

$$\psi_j : M^{(m-i,i)} \rightarrow M^{(m-j,j)}$$

by

$$\psi_j(X) = \sum_{Z \subseteq_j X} Z,$$

for $X \subseteq_i [m]$. In the notation of tabloids

$$\psi_j(\{t\}) = \sum_{\{s\}} \{s\},$$

where the sum is over all the $(m-j,j)$-tabloids $\{s\}$ whose second row is a subset of the second row of $\{t\}$.

Proposition 2. *Let t be an $(m-i,i)$-tableau and e_t^j its j-polytabloid. Let $k < j \le i$ then*

$$\psi_k(e_t^j) = 0,$$

while

$$\psi_j(e_t^j) = e_{t'}^j,$$

where t' is the $(m-j,j)$-tableau obtained from t by moving the last $i-j$ entries in the bottom row to the end of the top row.

Proof.

$$\psi_k(e_t^j) = \psi_k(\kappa_j(t)\{t\})$$
$$= \psi_k\left(\sum_{\sigma \in C_j(t)} (-1)^\sigma \sigma\{t\}\right)$$
$$= \sum_{\sigma \in C_j(t)} (-1)^\sigma \sigma \psi_k(\{t\})$$
$$= \sum_{\sigma \in C_j(t)} (-1)^\sigma \sigma \sum_{\{s\}} \{s\},$$

where the second sum is over all $(m-k,k)$-tabloids, $\{s\}$, whose second row is a subset of the second row of $\{t\}$. As $k < j$ then some element in the first j entries of the second row of t must lie in the top row of $\{s\}$, and there must be a transposition $\sigma \in C_j(t)$ fixing $\{s\}$. The terms involving this transposition and the terms not involving this transposition have opposite signs, and cancel, thus

$$\psi_k(e_t^j) = 0.$$

Similarly

$$\psi_j(e_t^j) = \sum_{\sigma \in C_j(t)} (-1)^\sigma \sigma \sum_{\{s\}} \{s\},$$

with cancellation for any $\{s\}$ for which one of the first j elements of the second row of t appears in the first row. Therefore

$$\psi_j(e_t^j) = \sum_{\sigma \in C_j(t)} (-1)^\sigma \sigma\{t'\}$$
$$= \sum_{\sigma \in C_j(t')} (-1)^\sigma \sigma\{t'\}$$
$$= e_{t'}^j.$$

$\square$

So, when restricted to $S^{(m-i,j)(m-i,i)} \subseteq M^{(m-i,i)}$ the image of ψ_j is isomorphic to $S^{(m-j,j)}$ and its kernel is $S^{(m-i,j)(m-i,i)}$. This gives an alternative characterisation of $S^{(m-i,j)(m-i,i)}$ as:

$$S^{(m-i,j)(m-i,i)} = \cap_{k=0}^{j-1}(\ker(\psi_k : M^{(m-i,i)} \to M^{(m-k,k)})).$$

To prove Theorem 1 in the next section we will identify a submodule of $M^{(m-i,i)}$ and study its images under this map. We shall conclude this section by stating a special case of the famous hook length formula [1], which gives the dimension of a Specht module.

Theorem 3. *Let k be a field.*

$$dim S^{(m-j,j)} = \binom{m}{j} - \binom{m}{j-1},$$

where $\binom{m}{-1} = 0$.

The astute reader will have noticed that is the term which appears in the formula in Theorem 1.

3. Proof of Theorem 1

The rows of the inclusion matrix $A_i^n(m)$ are indexed by the i subsets of $[m]$, which is naturally the basis for the $k\mathfrak{S}_m$-module $M^{(m-i,i)}$. The columns of span a submodule of $M^{(m-i,i)}$ of dimension $\operatorname{rank}_k(M^{(m-i,i)})$. We shall denote this submodule by $P_i^n(m)$. Our analysis of $M^{(m-i,i)}$ in the previous section gives rise to a chain of submodules

$$P_i^n(m) \supseteq P_i^n(m)_1 \supseteq P_i^n(m)_2 \supseteq \cdots \supset P_i^n(m)_i \supseteq 0,$$

where $P_i^n(m)_j := P_i^n(m) \cap S^{(m-i,j)(m-i,i)}$. This shows that

$$P_i^n(m) \sim \begin{matrix} L^{(m)} \\ L^{(m-1,1)} \\ \vdots \\ L^{(m-i,i)} \end{matrix},$$

where each $L^{(m-j,j)}$ is some submodule of $S^{(m-j,j)}$. In particular, $L^{(m-j,j)}$ is the image of $P_i^n(m)_j$ under the map $\psi_j : M^{(m-i,i)} \to M^{(m-j,j)}$.

Observe that the columns of the matrix $A_i^n(m)$ correspond to the images of the n-subsets of $[m]$ under the homomorphism $\psi_i : M^{(m-n,n)} \to M^{(m-i,i)}$. The module $P_i^n(m)$ is thus the image of $\psi_i : M^{(m-n,n)} \to M^{(m-i,i)}$. Observe that although $(m-n,n)$ may not be a partition, we can still define the permutation module $M^{(m-n,n)}$, and we can also define j-polytabloids for any $j < \min\{m-n,n\}$. Denote by $x \in P_i^n(m)$ the image of the j-polytabloid corresponding to a $(m-n,n)$-tableau t, for $j \leq i$. Our assumption from the introduction that $i < \min\{m-n,n\}$ ensures that this j-polytabloid is well-defined.

$$\begin{aligned} x :&= \psi_i(e_t^j) \\ &= \sum_{\sigma \in C_j(t)} (-1)^\sigma \sigma \sum_{\{s\}} \{s\} \end{aligned}$$

where the second sum is over all $(m-i,i)$-tabloids, $\{s\}$, whose second row is a subset of the second row of $\{t\}$. Of course we have cancellation of any terms for which the first j entries from the second row of t do not appear in the second row of $\{s\}$, so the sum is over all $(m-i,i)$-tabloids whose second row is a subset of the second row of $\{t\}$ of size i containing these first j entries. Observe then that

$$x = \sum_s e_s^j,$$

where the sum is over all $(m-i,i)$-tableaux obtained by moving $n-i$ of the last $n-j$ entries of the second row of t to the top row. As x is a sum of j-polytabloids, $x \in S^{(n-i,j)(m-i,i)}$ and thus $x \in P_i^n(m)_j$, and so its image under ψ_j is in $L^{(m-j,j)}$.

Proposition 4. *If* $p \nmid \binom{n-j}{i-j}$ *then* $L^{(m-j,j)} = S^{(m-j,j)}$.

Proof. Let $x = \sum_s e_s^j$ as above. Then, by Proposition 2,

$$\psi_j(x) = \psi_j(\sum_s e_s^j)$$

$$= \sum_s e_{s'}^j$$

where s' is the $(m-j, j)$-tableau obtained from s by moving the last $i-j$ entries of the bottom row to the top row. Each term is equal, and and the sum is over all $(m-i, i)$-tableaux obtained by moving $n-i$ of the last $n-j$ entries of the second row of t to the top row, of which there are $\binom{n-j}{i-j}$. Thus

$$\psi_j(x) = \binom{n-j}{i-j} e_{s'}^j,$$

and hence $e_{s'}^j \in L^{(m-j,j)}$. In fact, this shows that any j-polytabloid is in $L^{(m-j,j)}$ and thus $L^{(m-j,j)} = S^{(m-j,j)}$. $\qquad\square$

Proof of Theorem 1. Consider the image of $P_i^n(m)_j$ under the map ψ_j, which is a submodule $L^{(m-j,j)} \subseteq S^{(m-j,j)}$. By Proposition 4, if $p \nmid \binom{n-j}{i-j}$ then $L^{(m-j,j)} = S^{(m-j,j)}$. On the other hand, if $p \mid \binom{n-j}{i-j}$ then for any column $y = \psi_i(Y)$ of $A_i^n(m)$,

$$\psi_j(y) = \sum_{X \subset_i Y} \psi_j(X)$$

$$= \sum_{X \subseteq_i Y} \sum_{Z \subseteq_j X} Z$$

$$= \sum_{Z \subseteq_j Y} \binom{n-j}{i-j} Z$$

$$= 0.$$

This means that ψ_j is the zero map on $P_i^n(m)$, thus $L^{(m-j,j)} = 0$. We conclude that

$$P_i^n(m) \sim \begin{matrix} L^{(m)} \\ L^{(m-1,1)} \\ \vdots \\ L^{(m-i,i)} \end{matrix},$$

with

$$L^{(m-j,j)} = \begin{cases} 0 & \text{if } p \mid \binom{n-j}{i-j} \\ S^{(m-j,j)} & \text{if } p \nmid \binom{n-j}{i-j} \end{cases}.$$

The dimension of this module is then:

$$\dim_k(P_i^n(m)) = \sum_{j=0}^{m} \dim_k(L^{(m-j,j)})$$

$$= \sum_{p \nmid \binom{n-j}{i-j}} \dim_k(S^{(m-j,j)})$$

$$= \sum_{p \nmid \binom{n-j}{i-j}} \binom{m}{j} - \binom{m}{j-1},$$

where the last equality is due to Theorem 3. The rank of the inclusion matrix $A_i^n(m)$ over k is the dimension of the $k\mathfrak{S}_m$-module $P_i^n(m)$, thus proving the result. $\square$

Acknowledgements

References

[1] FRAME, J. S., ROBINSON, G. D. B., & Thrall, R. M. (1954). The hook graphs of the symmetric group. Canadian Journal of Mathematics, 6, 316-324.

[2] FRANKL, P. (1990). Intersection theorems and mod p rank of inclusion matrices. Journal of Combinatorial Theory, Series A, 54(1), 85-94.

[3] FRUMKING, A., & YAKIR, A. Rank of inclusion matrices and modular representation theory. Israel J. Math., 71(3):309–320, 199

[4] GOTTLIEB, D. H. (1966). A certain class of incidence matrices. Proceedings of the American Mathematical Society, 17(6), 1233-1237.

[5] JAMES, G. D. (2006). The representation theory of the symmetric groups (Vol. 682). Springer.

[6] JAMES, G. D. (1977). A characteristic-free approach to the representation theory of G_n. Journal of Algebra, 46(2), 430-450.

[7] LINIAL, N., & ROTHSCHILD, B. L. (1981). Incidence matrices of subsets—a rank formula. SIAM Journal on Algebraic Discrete Methods, 2(3), 333-340.

[8] WILSON, R. M. (1990). A diagonal form for the incidence matrices of t-subsets vs. k-subsets. European Journal of Combinatorics, 11(6), 609-615.

DEPARTMENT OF PURE MATHEMATICS AND MATHEMATICAL STATISTICS, UNIVERSITY OF CAMBRIDGE, CAMBRIDGE, CB3 0WA

Email address: ljj33@cam.ac.uk