Finding the Cardinality of a Grassmann Variety over a Finite Field

Nicholas Lucero
(nlucero@mit.edu)

March 16, 2005

Consider the field F_q, consisting of q elements. For any non-negative integer n, F_q^n is the n-dimensional vector space consisting of n-tuples of of elements in F_q. Also, for any integer k such that $0 \leq k \leq n$, the Grassmann variety $G(k,n)(F_q)$ is the set of all k-dimensional subspaces of F_q^n. In this presentation, we shall attempt to count the elements of $G(k,n)(F_q)$.

A few (more) definitions: A *basis* of a k-dimensional vector space V over F_q is a subset v_1, v_2,..., v_k of $\dim(V) \times 1$ column vectors in V that are linearly independent and span V. The condition of linear independence is met if

$$\text{For } c_i \in F_q, \qquad \sum_{i=1}^{k} c_i v_i = 0 \quad \Rightarrow \quad c_i = 0 \quad \forall i \tag{1}$$

The vectors v_i form a basis of V if and only if every $v \in V$ can be written as

$$v = \sum_{i=1}^{k} c_i v_i \tag{2}$$

for some choice of $c_i \in F_q$. Note that the choice of constants c_i must be unique for every v since there are exactly q^k elements of V and exactly q^k distinct choices for the c_i. Or, more rigorously, suppose there are two sets of constants c_i and d_i in F_q such that $\sum c_i v_i = \sum d_i v_i = v$. Then $\sum (c_i - d_i) v_i = v - v = 0$ and we see that $c_i = d_i$ for all i. Equation 2 also implies that the basis spans *only* V. That is, $\sum c_i v_i \in V$ for any choice of $c_i \in F_q$.

We will represent the set of basis vectors as the $k \times n$ matrix $\mathbf{M}$.

$$\mathbf{M} = \begin{pmatrix} v_1 & v_2 & \cdots & v_k \end{pmatrix}^T = \begin{pmatrix} v_1^T \\ v_2^T \\ \vdots \\ v_k^T \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & & \vdots \\ \vdots & & \ddots & \vdots \\ a_{k1} & \cdots & \cdots & a_{kn} \end{pmatrix} \tag{3}$$

for $a_{ij} \in F_q$. Thus, there is a unique $k \times 1$ vector $c_v \in F_q^k$ such that $\mathbf{M}^T c_v = v$. Our matrix representation provides a very handy way to look at things.

Proposition: If $\mathbf{M}$ is a matrix of basis vectors for the subspace V, the rows of any matrix resulting from row operations on $\mathbf{M}$ also form a basis for V.

Proof. Let $\mathbf{B}$ be a $k \times k$ matrix with entries in F_q. Since $\tilde{\mathbf{M}} = \mathbf{BM}$ is a matrix resulting from row operations on $\mathbf{M}$, we will show that $\tilde{\mathbf{M}}^T c_v \in V$ for all $c_v \in F_q^k$ and for any choice of $\mathbf{B}$.

$$\tilde{\mathbf{M}}^T c_v = (\mathbf{BM})^T c_v = \mathbf{M}^T \mathbf{B}^T c_v = \mathbf{M}^T c_{\tilde{v}} = \tilde{v} \in V \tag{4}$$

Since $c_{\tilde{v}}$ can be any element in F_q^k, we see that $\mathbf{M}$ and $\tilde{\mathbf{M}}$ are both basis matrices for the same subspace. □

Proposition: Every basis matrix $\mathbf{M}'$ of a subspace V is equal to $\mathbf{BM}$ for some $\mathbf{B}$ (as defined above), where $\mathbf{M}$ is another basis matrix for V.

Proof. Let the basis vectors composing $\mathbf{M}$ be v_i, as before. In the same manner, let the basis vectors composing $\mathbf{M}'$ be v_i'. Since all of the v_i' are necessarily in V, we know the following:

$$\mathbf{M}' = \begin{pmatrix} {v_1'}^T \\ {v_2'}^T \\ \vdots \\ {v_k'}^T \end{pmatrix} = \begin{pmatrix} \sum_{i=0}^{k} b_{1i} {v_i}^T \\ \sum_{i=0}^{k} b_{2i} {v_i}^T \\ \vdots \\ \sum_{i=0}^{k} b_{ki} {v_i}^T \end{pmatrix} \tag{5}$$

for some set of $b_{ij} \in F_q$. However,

$$\begin{pmatrix} b_{11}v_1^T + b_{12}v_2^T + \ldots + b_{1k}v_k^T \\ b_{21}v_1^T + b_{22}v_2^T + \ldots + b_{2k}v_k^T \\ \vdots \\ b_{k1}v_1^T + b_{k2}v_2^T + \ldots + b_{kk}v_k^T \end{pmatrix} = \begin{pmatrix} b_{11} & b_{12} & \cdots & b_{1k} \\ b_{21} & b_{22} & & \vdots \\ \vdots & & \ddots & \vdots \\ b_{k1} & \ldots & \ldots & b_{kk} \end{pmatrix} \begin{pmatrix} {v_1}^T \\ {v_2}^T \\ \vdots \\ {v_k}^T \end{pmatrix} = \mathbf{BM} \tag{6}$$

for some $\mathbf{B}$. □

We have shown that more than one matrix can be a basis matrix for the same subspace. It is important for us to ensure that we cast the basis matrices in a form that guarantees them to represent one distinct subspace each. This can be accomplished by using row operations to cast $\mathbf{M}$ in a reduced row echelon form $\mathbf{M}_R$, which insists the following:

- If a_{ij} is the first non-zero element in the i-th row of $\mathbf{M}_R$, then $a_{im} = \delta_{j,m}$.
- If a_{ij} and $a_{i'j'}$ are the first non zero elements in the i-th and i'-th rows, respectively, then $i > i' \Leftrightarrow j > j'$.

It is easy to see that any subspace V is represented by a unique reduced row echelon form matrix $\mathbf{M}_R$. Any other matrix representing V must be of the form $\mathbf{BM}_R$. However, $\mathbf{BM}_R$ is in reduced row echelon form if and only if $\mathbf{B}$ is the identity matrix.

Since every k-dimensional subspace of F_q^n can be generated by a set of k basis vectors v_i and any two bases which generate the same subspace have the same $\mathbf{M}_R$ representation, we may insist that there is an isomorphism between $k \times n$ reduced row echelon matrices with rank k and entries in F_q and k-dimensional subspaces of F_q^n. Note: basis matrices with k rows must have rank k if the basis vectors are to be linearly independent.

Definition: Let $M_R(k,n)(F_q)$ denote the set of all $k \times n$ reduced echelon matrices with entries in F_q. Thus $\#M_R(k,n)(F_q)$ is the number of such matrices.

Using the above definition, we may say that

$$\#M_R(k,n)(F_q) = \#G(k,n)(F_q) \tag{7}$$

We will count $M_R(k,n)(F_q)$ by considering an arbitrary element m in $M_R(k,n)(F_q)$. m has one and only one of the following properties:

Property 1: All entries in the first column of m are 0.

Property 2: The first column of m contains a single non-zero element. This element is 1 and resides in the first row.

Since any m has one of these properties and the properties are mutually exclusive, we know that $\#M_R(k,n)(F_q)$ can be obtained by adding the number of elements with property 1 to the number of elements with property 2. We will begin by counting the elements with property 1.

Any matrix $m \in M_R(k,n)(F_q)$ with property 1 can be drawn in the following manner:

$$m = \begin{pmatrix} 0 & \\ \vdots & m' \\ 0 & \end{pmatrix} \tag{8}$$

where m' is an element of $M_R(k,n-1)(F_q)$. Note that if $m' \notin M_R(k,n-1)(F_q)$ then $m \notin M_R(k,n)(F_q)$. Since there are $\#M_R(k,n-1)(F_q)$ elements in $M_R(k,n-1)(F_q)$, there are $\#M_R(k,n-1)(F_q)$ matrices in $M_R(k,n)(F_q)$ with property 1. Any matrix $m \in M_R(k,n)(F_q)$ with property 2 can be drawn in the following manner:

$$m = \left(\begin{array}{c|c} 1 & v' \\ \hline 0 & \\ \vdots & m' \\ 0 & \end{array}\right) \tag{9}$$

where $v' \in F_q^{n-1}$ and m' is a $(k-1) \times (n-1)$ matrix. (the bars within the matrix are present only to act as delimiters).

Fairly Straightforward Proposition: The matrix m' in equation 9 is an element of $M_R(k-1,n-1)(F_q)$.

Proof. The matrix m' is clearly composed of elements of F_q since m is composed of elements of F_q. Also, m' must be in reduced echelon form because we insist that m is in reduced echelon form. We must verify that m' is a basis matrix for a $k-1$-dimensional subspace of F_q^{n-1}. The most trivial property checks out, since m' clearly has $k-1$ rows and $n-1$ columns. All that is left is to verify that the rows of m' are linearly independent. If the rows r_i' of m' were not linearly independent, then there would be a non-trivial set of $c_i \in F_q$ such that $\sum c_i r_i' = 0$. Since m' is reduced row echelon, any row that is linearly dependent would be set to zero by row operations. If a row of m' is zero, then a row of m is zero. Since we have already asserted that $m \in M_R(k,n)(F_q)$, no row of m is zero. Thus, the rows of m' are linearly independent and $m' \in M_R(k-1,n-1)(F_q)$. □

The number of matrices in $M_R(k-1,n-1)(F_q)$ is, of course, $\#M_R(k-1,n-1)(F_q)$. However, we must decide how v' affects the multiplicity. We may be drawn to the immediate conclusion that each m' has a multiplicity of q^{n-1}, since v' is composed of $n-1$ seemingly arbitrary elements of F_q. This is not the case. Since m' has $k-1$ independent rows, it has a rank of $k-1$. Thus, m' has $k-1$ pivot columns and only $(n-1)-(k-1)=n-k$ free columns. Since m is in reduced row echelon form, any entry in the first row of m that is directly above a pivot column of m' is zero. Only the entries of v' that are directly above free columns of m' are left alone (and are thus arbitrary). This means that for any given m', v' only has $n-k$ free variables. For any given m', the multiplicity is only q^{n-k}. Thus, there are $q^{n-k}\#M_R(k-1,n-1)(F_q)$ matrices m with property 2.

Bringing it all together: We have shown that

$$\#M_R(k,n)(F_q) = \#M_R(k,n-1)(F_q) + q^{n-k}\#M_R(k-1,n-1)(F_q) \tag{10}$$

Incorporating equation 7, we arrive at the result

$$\#G(k,n)(F_q) = \#G(k,n-1)(F_q) + q^{n-k}\#G(k-1,n-1) \tag{11}$$

which is a recursive formula for counting $G(k,n)(F_q)$. Our recursion will eventually terminate at the base cases $\#G(0,n')(F_q)$ and $\#G(k',k')$, both of which are 1. This is so because the only zero-dimensional subspace of any vector space is the subspace consisting of only the zero vector and the only k'-dimensional subspace of any k'-dimensional vector space is exactly itself.

In class, Professor Vogan presented the following result without proof:

$$\#G(k,n)(F_q) = \left[\begin{array}{c} n \\ k \end{array} \right]_q. \tag{12}$$

I spoke with him and found that he came upon this result via a different method than the one presented here. So far I have not been able to close equation 11 to obtain the above form. If you have any ideas, please let me know.