Finite Dimensional Inner Product Spaces

Jason Kenyon

November 2022

Contents

Preface

Hello

Chapter 1

Vector Spaces

1.1 test

Theorem 1.1.1. *test*

Chapter 2

Linear Functions

Chapter 3

Linear Systems of Equations

3.1 Rank

Definition 3.1.1. An elementary operation on an $m \times n$ matrix A is

1. Interchanging any two rows or columns of A
2. Scaling each entry in a row or or column of A
3. Adding a multiple of one row or column to another row or column of A

An elementary matrix is the result of applying one of the above to the $n \times n$ identity matrix I_n

Theorem 3.1.1. *Suppose that B is the result of applying an elementary row operation to B. Then there exists an elementary matrix E such that $B = EA$. Furthermore, E is the matrix obtained by performing the same elementary row operation to I_n. Similarly if B is the result of applying an elementary column operation to A, then there exits an elementary matrix E such that $B = AE$, and E is the result of applying the same elementary column operation to I_m.*

The proof is a tedious verification of cases; the elementary matrices are defined precisely for this to work.

Definition 3.1.2. The rank of a matrix A is defined as the rank of the linear function $L_A = Ax$

Theorem 3.1.2. *Let* $T : V \to W$ *be linear and* $A = [T]_\beta^\gamma$. *Then* $\operatorname{rank}(T) = \operatorname{rank}(L_A)$

Proof. Consider the map $\phi_\beta : V \to \mathbb{F}^n$. That is, the function mapping a vector to its representation in coordinates. This is linear by definition and invertible as we know that any basis represents a vector uniquely as a linear combination of its elements. We have

$$L_A(\mathbb{F}^n) = L_A\phi_\beta(V) = \phi_\gamma(T(V))$$

. It follows that

$$\dim(\operatorname{im}(L_A)) = \dim(\operatorname{im}(T))$$

because ϕ_γ is an isomorphism. ■

Theorem 3.1.3. *Let* A *be* $m \times n$ *and* P, Q $m \times m$ *and* $n \times n$. *Then*

1. $\operatorname{rank}(AQ) = \operatorname{rank}(A)$
2. $\operatorname{rank}(PA) = \operatorname{rank}(A)$
3. $\operatorname{rank}(PAQ)$

Proof.

$$\begin{aligned}
\operatorname{im}(L_{AQ}) &= \operatorname{im}(L_A L_Q) && (3.1)\\
&= L_A L_Q(\mathbb{F}^n) && (3.2)\\
&= L_A(L_Q((\mathbb{F}^n)) && (3.3)\\
&= L_A(\mathbb{F}^n) && (3.4)\\
&= \operatorname{im}(L_A) && (3.5)
\end{aligned}$$

Thus, $\operatorname{rank}(L_{AQ}) = \operatorname{rank}(L_A)$. Similarly, $\operatorname{im}(L_P L_A) = L_P(\operatorname{im}(L_A)) = \operatorname{im}(L_A)$ and so $\dim(\operatorname{im}(L_P L_A)) = \dim(\operatorname{im}(L_A))$ since P is an isomorphism. It follows, by applying the previous two results that $\operatorname{rank}(PAQ) = \operatorname{rank}(A)$. ■

Theorem 3.1.4. *Let*

$$A = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{1m} & \cdots & a_{mn} \end{pmatrix}.$$

Then $\operatorname{rank}(A) = \dim\left(\operatorname{span}\left\{\begin{pmatrix} a_{11} \\ \vdots \\ a_{1m}\end{pmatrix}, \cdots \begin{pmatrix} a_{1n} \\ \vdots \\ a_{mn}\end{pmatrix}\right\}\right)$

Proof.

$$\begin{aligned} \operatorname{im}(L_A) &= L_A(\mathbb{F}^n) && (3.6)\\ &= L_A(\operatorname{span}\{e_1, \dots e_n\}) && (3.7)\\ &= \operatorname{span}\{Ae_1, \dots, Ae_n\} && (3.8)\\ &= \operatorname{span}\left\{ \begin{pmatrix} a_{11} \\ \vdots \\ a_{1m} \end{pmatrix}, \cdots \begin{pmatrix} a_{1n} \\ \vdots \\ a_{mn} \end{pmatrix} \right\} && (3.9) \end{aligned}$$

Furthermore, $\dim(\operatorname{span}(X))$ is nothing but the number of linearly independent vectors in X for any set of vectors X. Thus we have shown that the rank of a matrix is nothing but the number of linearly independent vectors in its columns. ■

Theorem 3.1.5. *Let A be an $m \times n$ matrix. Then a finite composition of elementary row and column operations applied to A results in a matrix of the form*

$$\begin{pmatrix} I_{\operatorname{rank}(A)} & O_1 \\ O_2 & O_3 \end{pmatrix}$$

where O_1, O_2, O_3 are zero matrices.

Proof. First, note that if A is a zero matrix, then by theorem 3.1.4 $\operatorname{rank}(A) = 0$, and so $A = I_0$, the degenerate case of our claim. Suppose otherwise. We proceed by induction on m, the number of rows of A. In the case that $m = 1$, we may convert A to a matrix of the form

$$\begin{pmatrix} 1 & 0 & \cdots & 0 \end{pmatrix}$$

by first making the leftmost entry 1 and adding the corresponding additive inverses of the others to the other columns. Clearly the rank of the above matrix is 1 and is of the form

$$\begin{pmatrix} I_1 & O \end{pmatrix}$$

This is another degenerate case, as it lacks zeros below the identity. Now suppose that our theorem holds when A has $m - 1$ rows.

To demonstrate that our theorem holds when A is an $m \times n$ matrix, notice that when $n = 1$, we can argue that our theorem holds as before, but using row operations instead of column operations. This is another degenerate case. For $n > 0$, note that there exists an entry $A_{ij} \neq 0$ and by applying at most an elementary row and column operation, we can move A_{ij} to position $1, 1$. Additionally, we may transform A_{ij} to value 1, and as before, transform all of the entries in row and column 1 besides A_{ij} to 0. Thus we have a matrix of the form

$$\begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & x_{11} & \cdots & x_{1\ n-1} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & x_{m-1\ 1} & \cdots & x_{m-1\ n-1} \end{pmatrix}$$

■

The submatrix defined by x_{ij} is of dimension $m - 1 \times n - 1$ and so must have rank $\text{rank}(A) - 1$ as elementary opertations preserve rank and deleting a row and column of a matrix reduces its rank by 1. Furthermore, by our induction hypothesis the above matrix may be converted via a finite number of elementary operations to a matrix of the form

$$\begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & I_{\text{rank}(A)-1} & O_1 & \\ \vdots & & & \\ 0 & O_2 & O_3 & \end{pmatrix}$$

Therefore, for an $m \times n$ matrix A, a finite number of elementary operations converts it into a matrix of the form

$$\begin{pmatrix} I_{\text{rank}(A)} & O_1 \\ O_2 & O_3 \end{pmatrix}$$

.

Theorem 3.1.6. *For any matrix A,* $\text{rank}(A^T) = \text{rank}(A)$*.*

Proof. By theorem 3.1.5, we may convert A to a matrix $D = BAC$ where $B = E_1 \cdots E_p$ and $C = G_1 \cdots G_q$ where E_i and G_i are elementary row and column matrices respectively. It follows that $D^T = C^T A^T B^T$, whence $\text{rank}(A^T) = \text{rank}(D^T)$ by theorem (insert) because elementary matrices are invertible, and so is the transpose of the compositions thereof. Further,

D^T must be of the same form as D since the only nonzero entries of D are along the diagonal from entry $1, 1$ to entry $\text{rank}(A), \text{rank}(A)$. Hence, we have $\text{rank}(A)$ linearly independent columns in the matrix D^T.

Since the columns of D^T are the rows of D, we see that the number of linearly independent columns of A is equal to the number of linearly independent columns of A^T. In other words, the dimension of the space generated by the columns of A is equal to the dimension of the space generated by its rows. ■

Theorem 3.1.7. *Let A be an invertible $n \times n$ matrix. Then A is a product of elementary matrices.*

Proof. By the dimension theorem, if A is invertible, then $\text{rank}(A) = n$. So by theorem 3.1.5 A may converted into a matrix of the form $I_n = E_1 \cdots E_p A G_1 \cdots G_q$, whence $A = E_1^{-1} \cdots E_p^{-1} I_n G_1^{-1} \cdots G_q^{-1}$. ■

Theorem 3.1.8. *Let $T : V \to W$ and $U : W \to Z$. Let A and B be matrices for which AB is defined. Then*

1. $\text{rank}(TU) \leq \text{rank}(U)$
2. $\text{rank}(TU) \leq \text{rank}(T)$
3. $\text{rank}(AB) \leq \text{rank}(A)$
4. $\text{rank}(AB) \leq \text{rank}(B)$

The proof follows from the fact that $\text{im}(UT) = UT(V) = U(T(V))) = U(\text{im}(T)) \subseteq U(W) = \text{im}(U)$ and the isomorphism between linear functions and matrices.

3.2 Form

We now apply the fruits of our investigation into vector spaces and linearity to solve systems of linear equations.

Definition 3.2.1. A linear system of equations is a collection of m equations of the form:

$$a_1 x_1 + \cdots + a_n x_n = b$$

where $a_i, x_i, b \in \mathbb{F}$ for $1 \leq i \leq n$. Equivalently, we may say $Ax = b$ for an $m \times n$ matrix A, where $x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ and $b = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix}$. If $b = \mathbf{0}$, the linear system is said to be homogenous.

Definition 3.2.2. A solution to a linear system is a vector $s \in \mathbb{F}^n$ such that $As = b$

Theorem 3.2.1. *Let A be an $m \times n$ matrix over $\mathbb{F}$. If $m < n$, then the homogenous system $Ax = 0$ has a nontrivial solution.*

Proof. Notice that, the solution set to the system $Ax = 0$ is $\ker(L_A)$, so by the dimension theorem, $\dim(\ker(A)) = n - \operatorname{rank}(L_A)$. Additionally, we know that $\operatorname{rank}(A)$ is nothing but the number of linearly independent vectors defined by its rows which certainly cannot exceed m. Therefore $\operatorname{rank}(A) \leq m < n$, in which case $n - \operatorname{rank}(A) = \dim(\ker(A)) > 0$, and so $\ker(A) \neq \{0\}$. ■

Theorem 3.2.2. *For any solution s to the linear system $Ax = b$,*

$$\{s + s_0 : As_0 = \mathbf{0}\}$$

is its solution set.

Proof. Suppose that $As = b$ and $As' = b$. Then $A(s' - s) = As' - As = b - b = 0$. It follows that $s + (s' - s) \in S$. Conversely, if $y \in S$, then $y = s + s'$, in which case $Ay = A(s + s') = As + As' = b + 0 = b$. That is, $Ay = b$. ■

Theorem 3.2.3. *Let $Ax = b$ for an $n \times n$ matrix A. If A is invertible, then the system has a single solution $A^{-1}b$. If the system has a single solution, then A is invertible.*

Proof. Suppose A is invertible. Then $A(A^{-1}b) = AA^{-1}(b) = b$. Furthermore, if $As = b$ for some $s \in \mathbb{F}^n$, then $A^{-1}(As) = A^{-1}b$ and so $s = A^{-1}b$. Next, suppose that the system has a unique solution s. Then by theorem 3.2.2, we know that the solution set $S = \{s + s_0 : As_0 = 0\}$. But this is only the case if $\ker(A) = \{0\}$, lest s not be unique. And so, by the dimension theorem, A is invertible. ■

Theorem 3.2.4. *The linear system $Ax = b$ has a nonempty solution set if and only if* $\operatorname{rank}(A) = \operatorname{rank}(A|b)$.

Proof. If the system has a solution, then $b \in \mathrm{im}(L_A)$. Additionally, $\mathrm{im}(L_A) = L_A(F^n)$ and $L_A(e_i) = Ae_i = \begin{pmatrix} a_{1i} \\ \vdots \\ a_{ni} \end{pmatrix}$. Therefore, since $L_A(\mathbb{F}^n) = \mathrm{span}\{Ae_1, \ldots Ae_n\}$, $\mathrm{im}(L_A) = \mathrm{span}\{A_1, \ldots A_n\}$, where A_i is the i^{th} column of A. Certainly, $b \in \mathrm{span}\{A_1, \ldots A_n\}$ if and only if $\mathrm{span}\{A_1, \ldots A_n\} = \mathrm{span}\{A_1, \ldots A_n, b\}$, which is to say $\dim(\mathrm{span}\{A_1, \ldots A_n\}) = \mathrm{im}(\mathrm{span}\{A_1, \ldots A_n, b\})$, or, $\mathrm{rank}(A) = \mathrm{rank}(A|b)$. ■

3.3 Solution

Definition 3.3.1. A matrix of the form

$$\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix}$$

is said to be in reduced echelon form if

1. $a_{ii} \neq 0$ implies that $a_{ij} = 1$
2. $a_{ij} \neq 1$ implies that $a_{ij} = 0$
3. $a_{ij} = 0$ for all $1 \leq j \leq n$ implies that $i < r$ for all nonzero rows $\begin{pmatrix} a_{r1} & \cdots & a_{rn} \end{pmatrix}$

Theorem 3.3.1. *Any matrix can be converted into reduced echelon form via a finite number of elementary row operations.*

The proof should be clear from the definitions of the three elementary. This form is of particular interest because reducing an augmented matrix is equivalent to solving a linear system of equations. We now have a procedure for solving arbitrary systems of linear equations. For example, we may now demonstrate that a set of vectors is linearly dependent by finding a nontrivial solution to a linear system of equations; similarly we may apply theorem 3.2.4 to demonstrate that a set of vectors is linearly dependent. In the following chapter, we will also see that computing the elements of an eigenspace is made possible by reducing a matrix.

Chapter 4

Eigenspaces

Chapter 5

Orthogonality

5.1 Inner Products

Hello

Definition 5.1.1. There exists

Appendix A

Determinants as Permutations

Hello