CS 435 : LINEAR OPTIMIZATION FALL 2008

Lecture 6: Row Rank. Convex Sets.

Lecturer: *Sundar Vishwanathan*
COMPUTER SCIENCE & ENGINEERING INDIAN INSTITUTE OF TECHNOLOGY, BOMBAY

1 Row Rank of a matrix

The space $\{x \ : \ Ax = \vec{0}\}$ is called the null space of the matrix A. Having already seen that the dimension of the null space of A is equal to the $n - k$ where k is the number of linearly independent columns of the matrix A, we now examine how this relates to the number of linearly independent rows of A.

Indeed, we will show that it is equal to $n - r$ where r is the number of linearly independent rows of A. This, in some sense, is not surprising. Unlike the column space, here both x and the rows of A, have the same number of coordinates. As vectors, both x and rows of A are part of R^n. Recall that two vectors are perpendicular if their dot product is zero. What we are interested in is the set of all vectors x which are perpendicular to all rows of A. If the row space is the space spanned by the rows of A then what we are looking for is the orthogonal complement of the row space of A. It is hence believable that the dimension of the orthogonal complement is $n - r$

One way of proving this is to use Gaussian Elimination which we do next. We have proved that Gaussian Elimination does not change the set $\{x \ : \ Ax = \vec{0}\}$. From what we have done so far, it is easy to infer that Gaussian Elimination does not change the number of linearly independent columns of A. Why?

We will next prove that Gaussian Elimination does not change the space spanned by the rows. This is to be expected, since, we know that it does not change the orthogonal complement! So, one way to prove this is to show that the orthogonal complement of the orthogonal complement of a subspace U of R^n is the subspace U itself. We will do it differently but the reader is encouraged to try that approach.

LEMMA 1 *Gaussian Elimination does not change the row space of a matrix.*

PROOF: Gaussian Elimination consists of two elementary operations. Exchanging two rows and multiplying a row with a scalar and adding it to another row. It is clear that exchanging two rows does not change the row space.

Suppose row A_i is replaced by $A_i + cA_j$. This is the only change in the two sets of vectors. We are done if we can show that every vector in the new space is in the old space and vice versa. The argument is very similar to the exchange lemma we did when we proved that bases have the same size. Note that the new vector $A_i + cA_j$ is also present in the old space. Hence every vector in the new space belongs to the old space. Also A_i belongs to this new space since we can get it by subtracting cA_j from the new ith row. Hence vectors in the old space are also present in the new space. Since the space has not changed, the dimension remains unchanged. □

Gaussian Elimination provides us with a tool to replace the old set of (row) vectors with a new set with the same span. The new set has a structure which lets us make certain inferences, especially connected with the rank, easily.

LEMMA 1 $\dim(\{x \ : \ Ax = \mathbf{0}\}) \geq n - r$

PROOF: The previous lemma assures us that applying Gaussian Elimination does not alter either the dimension of the row space. We already know that it does not change the dimension of the space $\{x \ : \ Ax = \mathbf{0}\}$. Why? So, to answer our question, we will consider the system after applying Gaussian

Elimination. After the Gaussian elimination, A looks like:

$$\begin{pmatrix} 0 & \dots & 01 & \dots & \dots & \dots & \\ 0 & \dots & \dots & 01 & \dots & & \\ & & & \vdots & & & \\ 0 & \dots & \dots & \dots & 01 & \dots & \\ 0 & & & \dots & & & 0 \\ & & & \vdots & & & \end{pmatrix}$$

Here the tth row, A_t, for $1 \leq t \leq r$, contains a 1 in the i_tth coordinate and $1 \leq i_1 < i_2 < \cdots < i_r$. For simplicity of presentation, we will assume that $i_t = t$.

Actually, as per our old view, we did not necessarily have a one in that position. A_{t,i_t} contained the first non-zero entry in row t after Gaussian Elimination. We know further that $i_s < i_{s+1}$ for each $1 \leq s \leq r-1$. We assume that this constant A_{t,i_t} is one for convenience. If necessary we can divide the equation by a suitable constant–this does not change any parameter of relevance. Also we can without loss of generality assume $i_t = t$. We can do this, for example, by renumbering the variables.

Now let x be any solution to the set of equations. Then it can be easily seen that we can assign any arbitrary values to $x_{r+1} \dots x_n$, and solve for $x_1 \dots x_r$ to find a solution to the set of equations. This is very similar to the column case and we indeed tread the same path hereon. In particular, we look at the $n-r$ vectors of the form $U_{i-r} = (x_1, x_2, \dots, x_n); \quad i = r+1, \dots, n$, where the ith coordinate of U_{i-r} is 1 and for $j > r, j \neq i$, the jth co-ordinate is 0.

Then the U_i's are of the form:

$$\begin{pmatrix} U_{1,1} \\ \vdots \\ U_{1,r} \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \quad \begin{pmatrix} U_{2,1} \\ \vdots \\ U_{2,r} \\ 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix} \quad \dots \quad \begin{pmatrix} U_{n-r,1} \\ \vdots \\ U_{n-r,r} \\ 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix}$$

Where $U_{i,j}$ is the jth coordinate of the vector U_i and are obtained as outlined above for $1 \leq j \leq r$. Clearly these $n-r$ vectors $U_1, \dots, U_{n-r}$ are linearly independent, as the last $(n-r)$ coordinates show. Also each U_i is a solution to $Ax = \vec{0}$ by construction. Thus we have $(n-r)$ linearly independent solutions to $Ax = \vec{0}$. This proves the lemma. □

Now, as before, we will prove that the $\dim(\{x \ : \ Ax = \vec{0}\}) = n - r$. We do this by showing that every other vector in the space $\{x \ : \ Ax = \vec{0}\}$ can be written as a linear combination of the U_is.

Theorem 2 $\dim(\{x \ : \ Ax = \mathbf{0}\}) = n - r$

Proof: Consider any x' such that $Ax' = \vec{0}$. We construct the following new vectors:

$$\tilde{x} = x'_{r+1} U_1 + \cdots + x'_n U_{n-r}$$
$$x'' = x' - \tilde{x}$$

As before we are done if we prove that $x'' = \vec{0}$ since then $x' = \tilde{x}$ which is expressible as a linear combination of the U_is.

By construction,

$$x''_i = 0, \ i = r+1, \dots, n. \tag{1}$$

We now observe that

$$\begin{aligned} Ax^{''} &= A(x' - \tilde{x}) \\ &= Ax' - \sum_{i=1}^{n-r} x^{'}_{r+i} AU_i \\ &= \vec{0} - \sum_{i=1}^{n-r} x^{'}_{r+i}\vec{0} \quad \text{because} \quad AU_i = \vec{0} \\ &= \vec{0} \end{aligned} \tag{2}$$

Now we prove that $x^{''}_i = 0$ in the order $r \geq i \geq 1$. Consider the ith equation. Write this down to make sure you understand what we are talking about. The value of $x^{''}_i$ depends only on $x^{''}_j$ for $j > i$. Hence if all the co-ordinates $x^{''}_j$ for $j > i$ are zero then so is $x^{''}_i$. Whence, $x^{''} = 0$.

Thus we get $\dim(\{x \ : \ Ax = \vec{0}\}) = n - r$ □

This concludes the proof of the fact that the row rank of a matrix is equal to $n - t$ where t is the dimension of its null space. As we have already proved, this is also equal to the column rank of a matrix. We then get, as a corollary, that the row rank and column rank of a matrix are equal.

COROLLARY 3 *The row rank of a matrix is equal to its column rank.*

2 Solutions to $Ax = \mathbf{b}$

Consider now $Ax = \mathbf{b}$ for any arbitrary column vector $\mathbf{b}$. We see below that the set $\{x \ : \ Ax = \mathbf{b}\}$ looks like a subspace shifted by a vector. By shifting a subspace we mean take all points in a subspace and add a fixed vector to all of them.

THEOREM 4 *Let x_0 be a vector in $\mathbb{R}^n$ such that $Ax_0 = \mathbf{b}$. Then every solution to $Ax = \mathbf{b}$ can be written in the form $x_0 + x'$, where x' is any vector satisfying $Ax' = \vec{0}$.*

PROOF: Consider any $\tilde{x}$ such that $A\tilde{x} = \mathbf{b}$. Then, we have $Ax_0 = \mathbf{b}$ and $A\tilde{x} = \mathbf{b}$. Which means $A(\tilde{x} - x_0) = \vec{0}$. Set $x' = \tilde{x} - x_0$ to finish the proof. □

To conclude, the solution set of $Ax = \mathbf{b}$ looks like a subspace shifted by a vector x_0.

3 Convex Sets

We look at some of the geometric properties of sets of points in this section. Consider any two points $v_1$and v_2. Then the vector $\lambda v_1 + (1 - \lambda)v_2$ for $\lambda \in [0, 1]$, lies on the line segment joining v_1 and v_2. Indeed, as λ varies between 0 and 1 this vector covers all points in the segment between v_1 and v_2. When $\lambda = 0$ it is at v_2 and as λ increases, it moves continuously along the segment to end at v_1 when $\lambda = 1$.

Suppose you have three points v_1, v_2, v_3. These define a triangle. How do you describe all points lying inside this triangle?

If we consider a set of n points $S = \{v_1, \ldots, v_n\}$, then any point lying inside the object with $v_1, \ldots, v_n$ as its vertices can be written as $\sum_{i=1}^{n} \lambda_i v_i$, where $\sum_{i=1}^{n} \lambda_i = 1$ and $0 \leq \lambda_i \leq 1$. This is akin to taking linear combinations with more restrictions on the coefficients and is called a *convex combination.*

DEFINITION 1 *Given n vectors $v_1, \ldots, v_n$,*

$$\sum_{i=1}^{n} \lambda_i v_i; \quad 0 \leq \lambda_i \leq 1, \quad \sum_{i=1}^{n} \lambda_i = 1$$

is called a convex combination *of $v_1, \ldots, v_n$.*

Definition 2 *A set of points S is called* convex *if for any subset S' of S and for any point p which we get by convex combination of points in S', $p \in S$.*

This definition is unwieldy. Fortunately, there is an equivalent definition that we will use for this course. This essentially states that it is enough to restrict ourselves to sets S' of size two.

Definition 3 *A set of points S is called* convex *if for any two points u and v in S the segment joining the points u and v is in S'.*

Our interest in this definition stems from the fact that the set $\{x \ : \ Ax \leq \mathbf{b}\}$ is convex. This is because for any x, y satisfying $Ax \leq \mathbf{b}$ and $Ay \leq \mathbf{b}$, $A(\lambda x + (1 - \lambda)y) = \lambda Ax + (1 - \lambda)Ay \leq \lambda \mathbf{b} + (1 - \lambda)\mathbf{b} = \mathbf{b}$. Note the use of the fact that $0 \leq \lambda \leq 1$.