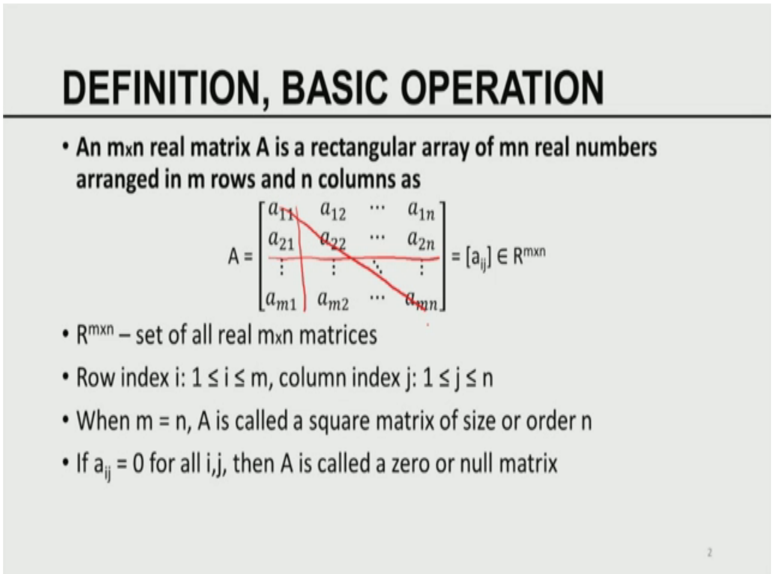
Dynamic Data Assimilation
Prof. S Lakshmivarahan
School of Computer Science
Indian Institute of Technology, Madras

Lecture – 05
Matrices

The last module 2.1, we reviewed the fundamental principles of vector spaces and various concept associated with vector space. In this module, 2.2, I am going to provide a quick overview of matrices. I am sure many of you have been introduced to various properties of matrices. I am going to collect all the properties that we would need at one place. So, I would like to make this module as a one stop shop where you can go back and refer to all the basic principles needed to pursue most of what we have to do in data assimilation.

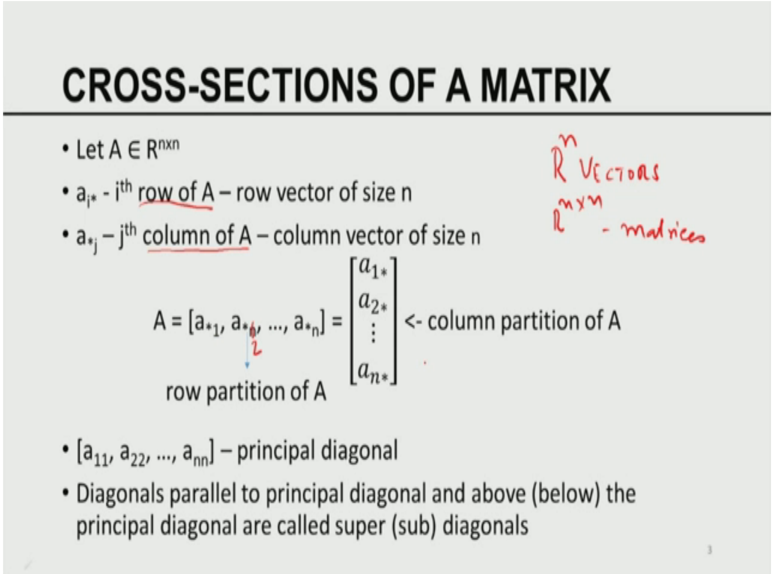
(Refer Slide Time: 00:50)



First definition and basic operations of matrices; a matrix m by n matrix is a real matrix if it has the n row or it has m rows and n columns. There are n m elements. Each row is a m vector. I am sorry each column is a m vector each row is the n vector if m is not equal to n. It is called a rectangular matrix i is. So, any typical element is called a i j i is called the row index j is called the column index when m is equal to n.

It is called a square matrix. Square matrixes of order n are size n order and size are used synonymously, if all the elements are 0. It is called a 0 matrix or a null matrix. We need a number 0. We need a null vector 0. We also need a null matrix 0. So, we will use the symbol 0, but the context will tell us whether we are talking about the number 0 or the vector 0 or the null matrix 0, but we need all these all these all these objects More often than not we will be dealing with square matrices.

(Refer Slide Time: 02:10)



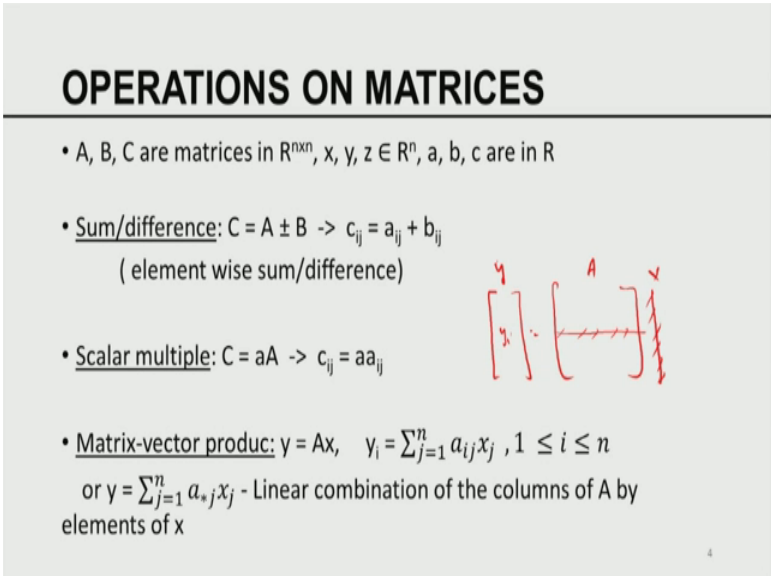
So, A belongs to R of m by n. I would like to be able to say a word or 2 about R n by n. Please recall, we have used the symbol R n to denote the set of all vectors. This is the set of all vectors. This is the set of all vectors likewise R n by n is the set of all matrices. So, there are n elements in a vector. There are m square elements in a matrix each element is a real number. So, there are infinitely many vectors there are infinitely many matrices in this set. So, I would like to emphasize this these sets R n. R n by n cross n; they are all infinite sets each one is a different object a vector is an object matrix as an object and so on.

I can refer to ith row or the jth column. So, yeah if you go back to the previous slide this is this is called the; this is called the first column. This is called the second row. So, we can talk about the notion of a row of a matrix. It is important to recognize row of a matrix a column of a matrix.

So, a matrix can be represented by represented by sequence of columns or a sequence of rows. So, the ith row is represented by a i star the jth column is represented by a star j. So, a star 1 is the first column a star 2. So, this must be this must be 2 here this is not n, I am sorry; this must be 2. So, a star 1, a star 2, a star n; these are n columns. A 1 star, a 2 star a and star these are n rows. So, this is called column partitioning this is called row partitioning. So, we can talk about partitions of a matrix.

Again going back to the previous slide the element that lie along the diagonal for example, in here that is called the diagonal of the matrix. So, column row diagonal these are called different cross sections of the matrix. So, a 1 1, a 2 2, a 3 is a vector of size n the vector that lies along the diagonal. So, that diagonal is called. So, called principal diagonal of a matrix all the diagonals that are parallel to the principal diagonal are called super diagonal or sub diagonal super diagonals are above the principal diagonal sub diagonals are below the principal diagonal.

(Refer Slide Time: 05:17)



These are all nomenclature that one has to remember these are fundamental to our pursuit of mathematical treatment of data simulation.

Now, I am going to define quickly several operations and matrices, I would like to back up a little and then explain this. Now if I have set of integers, I need to define operations integer addition multiplication subtraction. If you have real numbers you talk about addition multiplication subtraction division if you have vectors you talk about addition scalar product outer product and so on.

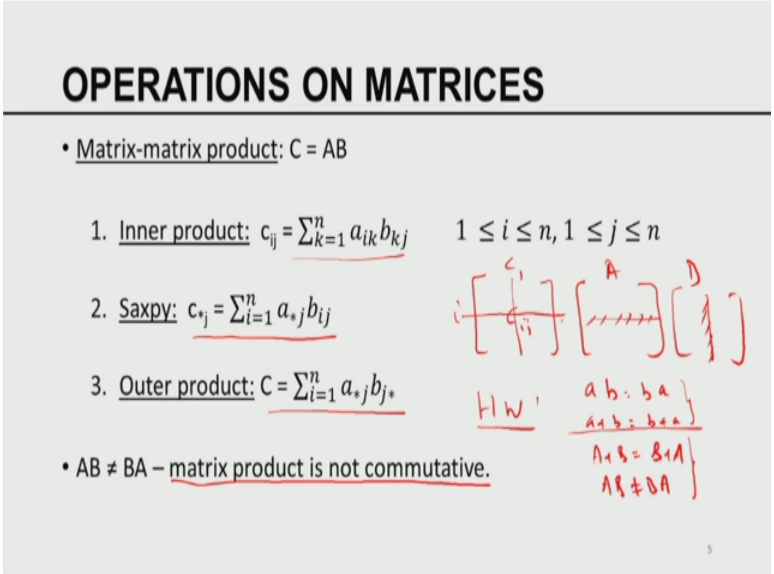
So, what does this tell you if you have a set of mathematical objects we have to talk about your sets of consistent operation for those sets? So, for number there are operations there real numbers operations recompensed number there are operations vectors that operation polynomial there are operations.

So, with respect to matrices likewise we have to have different sets of operations, I am going to quickly define some of the fundamental operations on matrices. So, if I define 3 matrices a b c, if I define 3 vectors x y z, if I define 3 numbers a b c; now look at this; now I have elements from 3 different animal kingdom matrices is one class of animals vectors is another class of animal scalars are another class of animals I am going to combine all of them to be able to do what I want to do. This is where the notion of a vector space comes into play some on difference of matrices is the matrix. So, c is the matrix which is the sum of a plus b; c is the matrix which is the difference of a minus b these sums are called element wise sum element wise difference.

If a is a matrix, a is a little a is a scalar, I can define a to be a times a; that is called scalar multiplication of a matrix you multiply the element each element of the matrix a by the scalar A. I also can combine matrix and vectors this is called matrix vector multiplication, I can define a vector y as the product of the matrix a and a vector x and that is defined by y i ith component of y is given by the ith the row of a times the vector. So, in here, I am going to represent a little picture to tell. So, if this is the vector y if this is the matrix a and this is the vector x to compute y i i take the ith row of a i multiply that by the vector x and that is the scalar product the ith row of a is given by a i j; j running from 1 to n.

The elements the vector x is given by x j j running from 1 to n. So, I am multiplying the first element with the first element second element with the second element; nth element with nth element summing in the map. So, it is the scalar product of the ith row and the vector x is the yth element, you continue to those for every one of them and that defines the vector that is what is called matrix vector product.

(Refer Slide Time: 08:26)



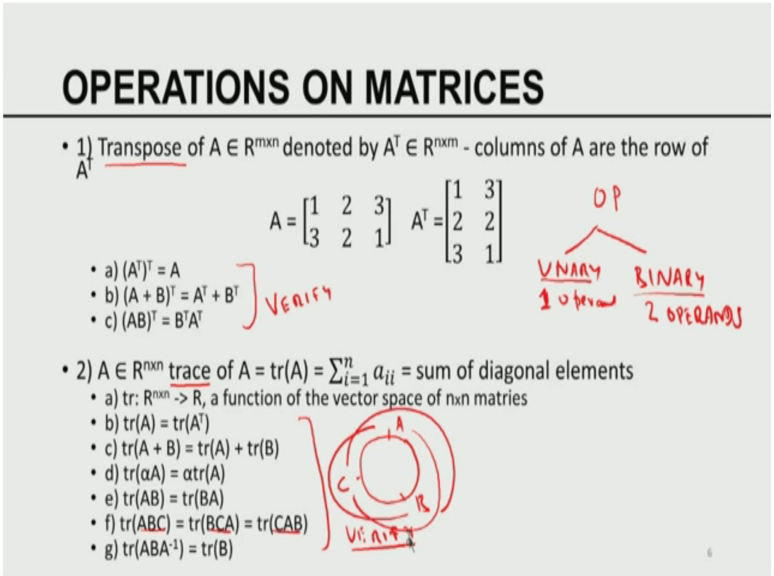
So, I can define now matrix; matrix product, we talked about several different operation previously look at this. Now sum of matrices difference of matrices multiplication of matrix by a constant multiplication of a matrix by a vector, now I am going to talk about multiplication of a matrix by a matrix multiplication of a matrix by a matrix is also a matrix it is given by the ith element of the matrix again, we all should know if I have a matrix c, if I have a matrix a, if I have a matrix b, this is a, this is b, if you consider the i jth element in here, this is the ith row, this is a jth column, this is the element c i j, the c i j is essentially the inner product of the ith times the jth column of b.

So, ith a row and j ith row of a and jth column of b the inner product is c i j that is given by this product, there are other ways of looking at the matrix product, one is called the Saxpy way another is called the outer product way, I have given these definitions in these I would like you to verify that the matrix product defined by the inner product Saxpy outer product they all give rise to the same result and I would like to be able to give that is a homework problem for you to work out; I think it will be a an illuminating homework for you to verify the matrix product can be defined in one of 3 ways.

I would like to now emphasize the fundamental property of matrix product matrix product is not commutative; that means, a b is not equal to b a let us go back now if you take 2 numbers a b, a b is equal to b a if you took 2 numbers a plus b is equal to b plus a, if you took 2 matrices a plus b is equal to b plus a, but if you take 2 matrices a b in general is not equal to b a.

So, what does this mean algebra of real numbers is commutative algebra matrix algebra is non commutative matrix product is not communicative and that is a very fundamental restriction when you go from real algebra to matrix algebra that one has to be cognizant off.

(Refer Slide Time: 10:48)



Now, I am going to define lots of other operations there are ten of operations and matrices which is very rich if I have a matrix which is m by n, I can define a matrix called A transpose of A which is denoted by A to the power T, it is the n by m. The rows of a or columns of A transpose and vice versa. So, if A is; this A transpose is that. So, what are the properties of transposes transpose operation transpose is a very fundamental and a basic operation.

So, transpose is called a urinary operation. So, I would like to now distinguish between 2 types of operation, operation can be either a binary operation a binary operation needs 2 operands for example, to add I need 2 numbers to multiply any 2 numbers to divide I need 2 numbers.

So, a biginary operation needs 2 operands a urinary operation on other hand needs only one operand what are the examples of urinary operation transpose; transpose of a negative of A inverse of A. So, transpose negative inverse they are all unity urinary operation addition subtraction multiplication they are all binary operation.

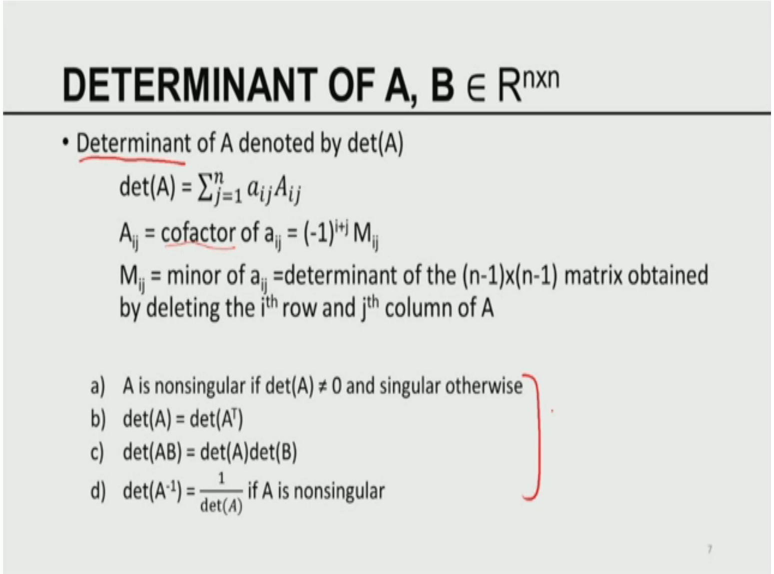
So, I would like you to be able to be cognizant of the fundamental difference between 2 types of operators binary operator binary operation urinary operator urinary operation this urinary operation transpose has several properties transpose of A transpose is itself transpose of the sum is the sum of the transposes transpose of the product is the product of the transposes these are all basic properties I am not going to prove them many books that I talked about at the end of module 2.1 has proofs of these in your case, if you do not want to prove this at least you should be able to verify how do you verify these take 2 matrices A and B, take a matrix A, do these operations and verify. So, I would like you to very strongly recommend please verify these properties is very fundamental to see why and how they operate they work.

The next urinary operation is called the trace of a matrix trace of a matrix is defined to be the sum of the elements of the diagonal. So, if I have a matrix a it is simply the sum of a i i when i is equal to one i get a 1 1, a 2 2, a 3 3, a n n. So, trace is a functional is a function from R into R you can think of it as a functional the trace has lots of important properties trace of a is equal to trace of A transpose I am assuming a is a n by n matrix trace of a plus b is trace a space of b trace of alpha times a is alpha time space of b trace of a b is equal to trace of b a trace is a same.

When you compute the product a b and b a trace of the product a b c is b c a and c b a, you can think of it as a circular property. So, this is a, this is b, this is c. So, you can think of a b c you can think of b c a, you can think of c a b you can you can run around the circle starting at a starting at b are starting at c. So, this; assume the property f essentially tells you no matter where you start the triple product have the same trace.

Then the trace of a times b times A inverse is simply a trace of b that that essentially comes from applying the property f to g and again I am going to leave all these things as a homework problem. I would like you to verify in other words these are simply definitions I would like you to be able to verify using simple examples is absolutely essential that we all have a good understanding of these properties

(Refer Slide Time: 15:11)

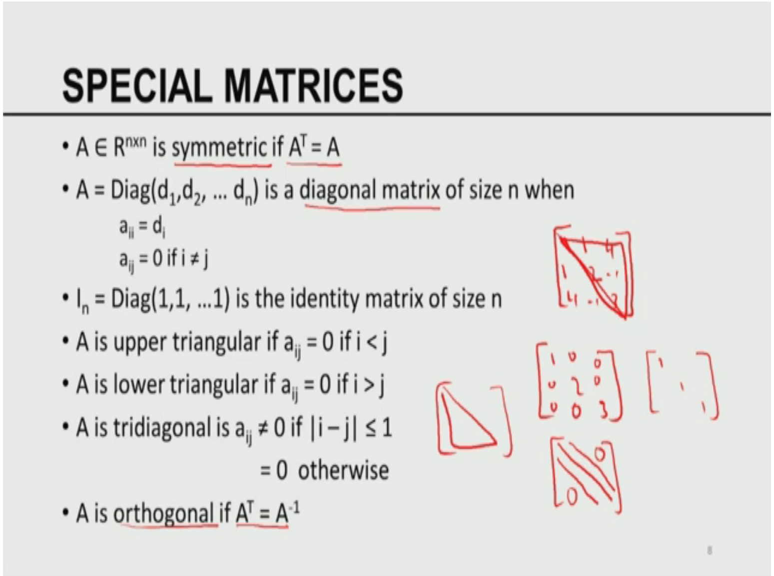


Then the notion of a determinant of a matrix, I am trying to list all the properties that the matrix process determinant of a matrix, we all know determinant is again is a function that maps a to real the determinant of a matrix is a number the determinant is defined by the product of the sum of the product of a i j with the cofactors everybody should have known the definition of a cofactor co-factor is called the signed minor. So, the determinant of a matrix is of fundamental quantity I am sure most of you should have been introduced to the notion of a determinant.

Now, I am going to introduce some of the properties of determinants if a is not singular determinant of a is not 0, if the determinant of a is 0, then the matrix is called singular determinant of a is equal to the determinant of A transpose determinant of A B is determinant of A times determinant of B determinant of A inverse is 1 over determinant of A; if a is non singular.

Again, these are the properties I am going to ask you to verify. So, what is the first thing ultimately you know how to prove, but the first step towards proving is to verify at least you should be confident to the fact yes this properties hold; I have already verified using examples, but examples verification is not a proof; proof is little bit more abstract a proof deals with all the cases verification deals only with specific instances. So, that is a difference between verification and proving is the ultimate goal, but to get to prove you need to verify first.

(Refer Slide Time: 16:58)



So, you need to build your expertise first to verify and then to prove. Now, I am going to enlist properties of several special matrices first of the property is called the symmetry A matrix A is said to be symmetric if A transposes A. So, what does this mean if I have a matrix if I have 1, 2, 3, if I have a 1 here, there must be a 1 here, if I have a 4 here that must be a 4 here, if I have a minus 1 here, I have a minus 1 here.

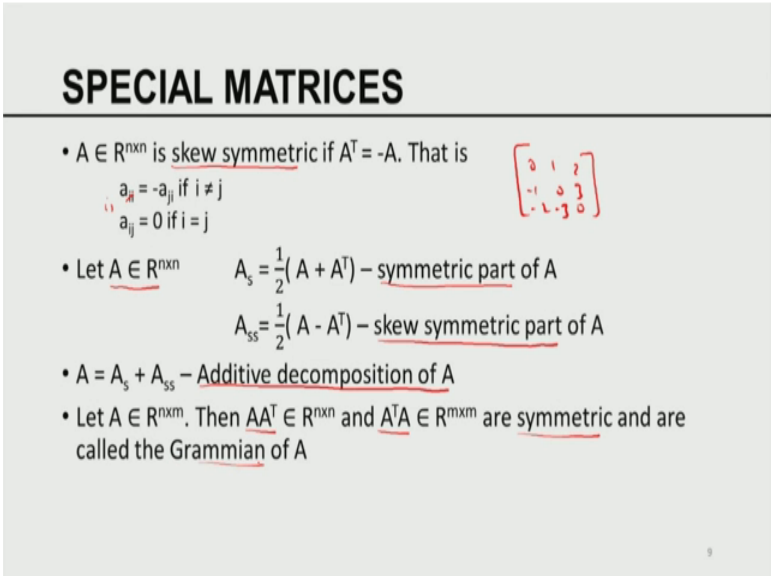
So, if I took the diagonal element the upper triangular part of the lower triangular part are mirror images of each other and that is what A transpose a refers to if A transpose is equal to a means the upper half and lower half are a mirror images of each other. So, symmetric matrices are special class of matrices there is a restriction the restriction comes from the fact the upper half must be a mirror image of the lower half a matrix could be a diagonal matrix in case there are only diagonal elements all the non diagonal elements are 0. So, what is an example of a diagonal matrix an example of diagonal matrix is 1 2 3, again 0 0 0 0 0 0; that is an example of a; that is an example of a diagonal matrix.

The unit matrix is a special kind of a diagonal matrix where all the elements along the diagonal are 1 1 1.

So, in this case, this is the diagonal matrix of size 3, this is also a diagonal matrices of size 3, this is a; this is a non unit matrix, this is a unit matrix, then we can talk about upper triangular matrix upper triangular matrices are those the diagonal and above the diagonal are non-zero. Lower triangular matrix is are those where the diagonal and the below the diagonal are all non-zero this is lower triangular that is upper triangular then a matrix can be tridiagonal; a tridiagonal matrix is one where the principal diagonal is non-zero the first super diagonal is non-zero. The first sub diagonal is not 0. Everybody else is 0; that is called tridiagonal matrix.

A matrix is said to be orthogonal matrix if A transpose is equal to a inverse. So, orthogonal matrix have this extremely nice special property that inverse is equal to the transpose; these are extremely basic pro these are examples of basic properties of special class of matrices.

(Refer Slide Time: 19:24)



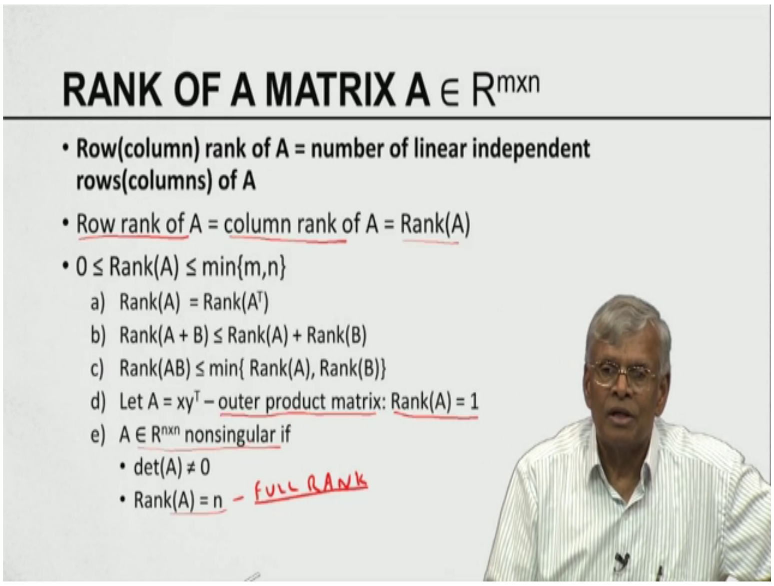
The special matrices continue a matrix is said to be skew symmetric matrix if A transpose is minus A. So, symmetric matrix is A transpose a skew symmetric matrices a transposes minus a. So, in a skew symmetric matrix a i; a i i in a skew symmetric matrices a i i is equal is equal to a I am sorry here it must be i j a i j must be equal to minus a j i and a i i i is equal to 0 if i is equal to j.

So, the diagonal elements of a symmetric matrix, skew symmetric matrix is 0. So, what is an example of a skew symmetric matrix 0 0 0 1 minus 1 2 minus 2 3 minus 3 that is an example of a skew symmetric matrix the diagonal elements are 0 there is a reflection, but the sign change. So, a i j is minus a j i given any matrix a i can separate the matrix into 2 parts one is called the symmetric part of the a another is called the skew symmetric part of a the symmetric part is one half of a and a a plus A transpose the skew symmetric part is one half of a minus A transpose.

Therefore, you can easily verify a is equal to a s plus symmetric part plus non symmetric part this is called the additive decomposition of a every general matrix can be expressed as a sum of a symmetric matrix consists of a symmetric part and a skew symmetric matrix consisting a skew symmetric part.

Given any matrix A; the product A transpose; the product A transpose A; these are 2 matrices; one can generate out of matrix A. In other words; given a matrix A compute A transpose. So, I have a and A transpose, I can multiply A and A transpose like this or A transpose A; A like that it turns out A; A transpose and A transpose A, both of them are symmetric and they have a special name they are called Grammian matrices gram is the one who first introduced this. So, these are called grammian of a Grammian of a are always symmetric for whatever a is.

(Refer Slide Time: 21:49)

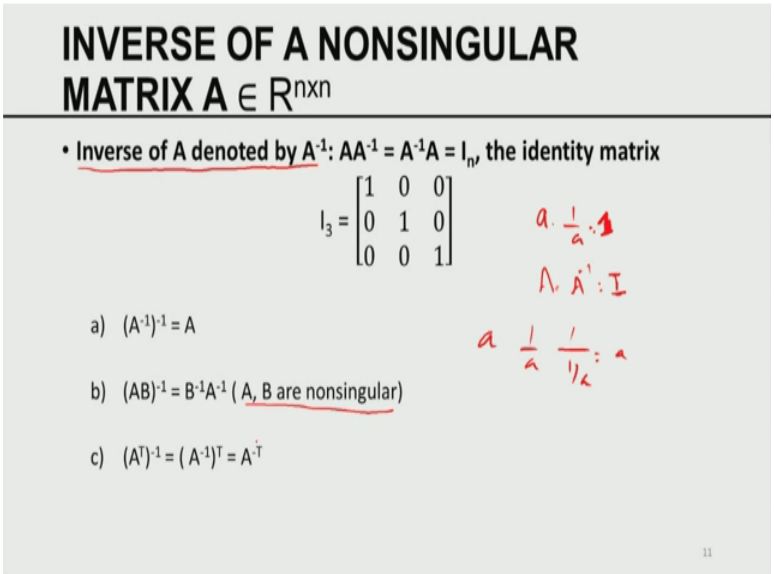


The next one is called the concept of rank of a matrix the rank of a matrix is essentially the number of linearly independent rows or columns. It can be shown the column rank is equal to the row rank. So, you can think of number of independent rows of a matrix called a row rank the number of yeah the number of independent rows of a matrix is called the row rank because the rows are vectors, if I have a bunch of vectors.

I can talk about the linear independence of a set of vectors the number of linearly independent vectors is called the rank. So, I can think of a row rank of a column likewise column rank of A the row of A is equal to column rank of A and the common value is called the rank of A. So, A is m by n matrix the rank of A is less than or equal to the minimum of m and n rank of A is rank is equal to A transpose of is the rank of A transpose rank of A sum is less than the sum of the ranks rank of a product is minimum of the rank of A and rank of B.

We have earlier seen outer product of 2 vectors is the matrix, if a matrix arises outer product of 2 vectors that matrix is always as a rank one if a matrix is non singular the rank of A is n. So, if A is n by n matrix; if it is non singular the determinant is not equal to 0, it is also the fact that the matrix is also n. So, you characterize the set of all non singular matrices to be those; the non-zero determinant or full rank. So, this is called the full rank; the full rank condition is same as non singularity is the same as determinant of a to be not equal to 0.

(Refer Slide Time: 24:15)



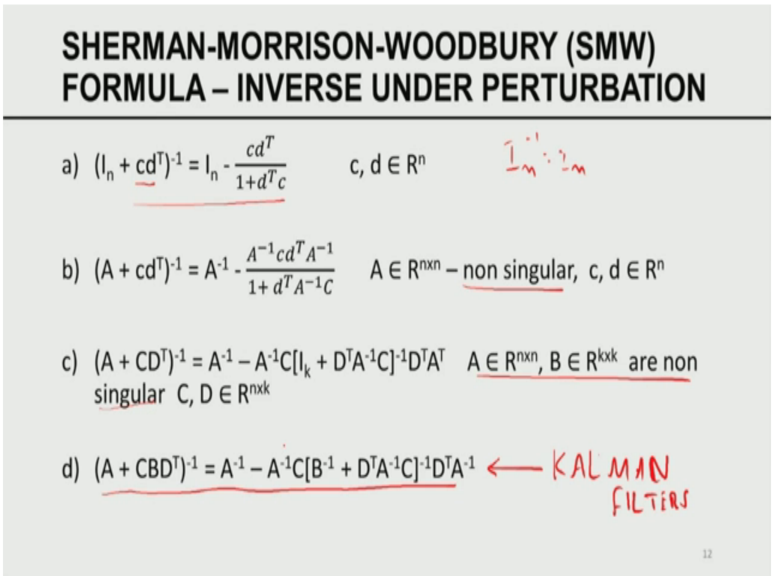
I am going to review the concept of inverse of a matrix inverse of A inverse of a matrix inverse of a matrix is denoted by A inverse and how do I define A inverse the same way in number theory. We say A times 1 over A is 1. We call 1 over A as the reciprocal in matrix theory. We call it A times A inverse is equal to I; we call it inverse.

So, inverse and reciprocal are pretty synonymous the role of number one in numbers is same as the role of number the matrix i in matrices, they are called unit elements if you multiply any number by 1 is same. If you multiply any matrix by identity is also one. So, what is one for numbers is identity matrix for matrices. So, I would like you to be able to know the property of an inverse matrix and through the inverse a times A inverse is i inverse of the urinary operation inverse of A inverse is a and that should not be surprising because reciprocal of a reciprocal is a given number.

So, I have a, I have 1 over a, that is equal to A. So, the reciprocal of reciprocal is a same number. So, inverse of inverse is A inverse of the product is the product of the inverses taken in the reverse order assuming the matrices are non singular these are fundamental property which will be used repeatedly in data assimilation inverse of A transpose is the transpose of the inverse and is a combined. So, inverse is one urinary operation transpose is another urinary operation; I am talking about the conjunction between conjunction of 2 urinary operations.

So, transpose is the inverse is the inverse of the transpose and that is denoted by A to the power minus T. So, A to the power of minus T means is A transpose inverse. I can perform any operation first any other operation second I can transpose an inverse or inverse and transpose both are same they are commutative.

(Refer Slide Time: 26:42)



Once you have the concept of inverse, I am going to introduce you to several different formulas relating to inverses these are called Sherman Morrison formula, these are called inverse under perturbation. So, if I have an identity matrix; the inverse of an identity matrix is equal to itself that is a general property, we all know one the reciprocal of one is one; there is a protocol of identity matrix is identity matrix, but if I add an outer product matrix to an identity matrix, it is no more an identity matrix, I know the inverse of identity matrix is its is identity.

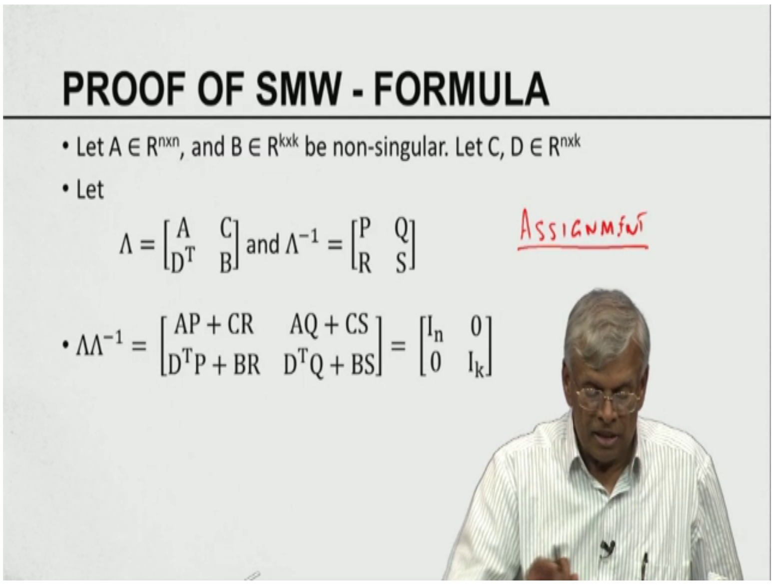
So, the question is this; if I perturb the identity matrix by an outer product matrix, we have already seen outer product matrix as a rank one the inverse is the sum can be expressed by this formula. This formula is called Sherman Morrison Woodbury formula.

So, c is the vector d is the vector c d transpose is an outer plan is an outer product matrix. So, this is called a rank one perturbation of I of n. So, if I perturb the matrix and compute the inverse I do not have to the compute the inverse from ground up I can simply update the inverse of I with this correction and that formula carries over and these are generalization I can be replace by a c and d remains the same now a remains the same a is non singular in this case also a is non singular in this case I am assuming a and b are non singular. I want you to be able to look at this.

Now c and d are vectors c and d are matrices. So, this is the most general form of this inverse operation this is the simple form of inverse operation it is this version that we will use repeatedly in Kalman filtering techniques. So, Sherman Meris Morrison Woodbury formula d is one of the most fundamental relation that is used in data assimilation especially in the derivation of Kalman filters; I am sorry Kalman filters.

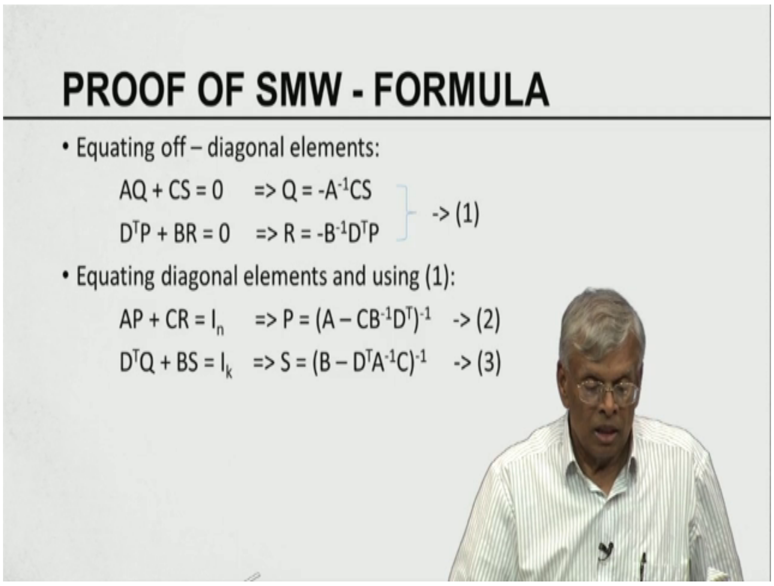
So, what does it tell; here if you know a matrix a is non singular, if you add a correction to that I can compute the inverse of the correction by simply a correction term the inverse of the original matrix. So, that is a very beautiful formula these formulas have been known since 1930s and mathematicians have done these things just for the fun of it and these formulas find great use in many of the derivations especially the ones relating to Kalman filters.

(Refer Slide Time: 29:44)



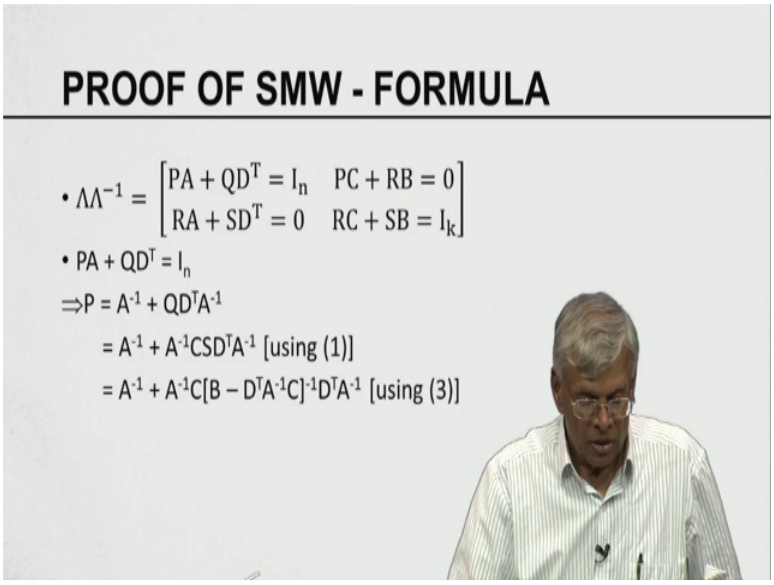
I have now given a proof of the Sherman Morrison Woodbury formula look at this. Now D is the generalized version of the Sherman Morrison Woodbury formula; I have given the proof of this, I am not going to go over the proof because it is given in extremely simple case. So, I am going to leave the proof as the reading assignment i would like you to be able to read the proof as an assignment.

(Refer Slide Time: 30:23)



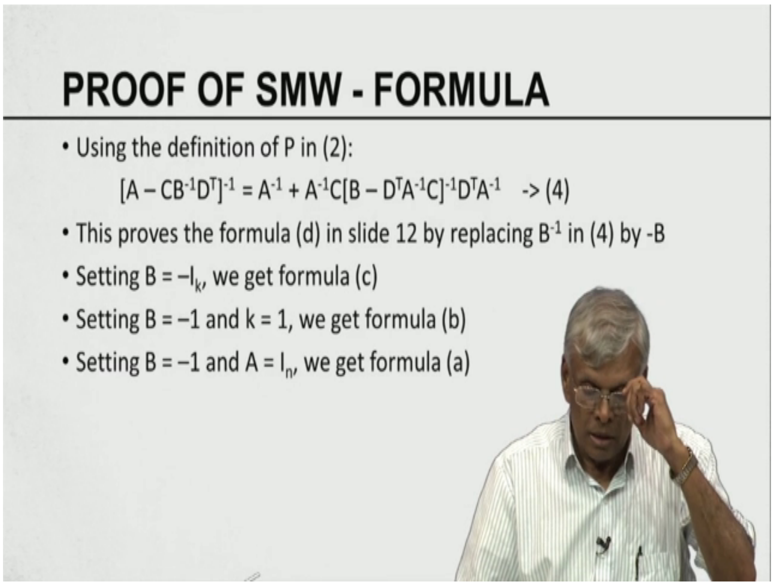
The proof is a reading assignment. I have given in extreme details.

(Refer Slide Time: 30:29)



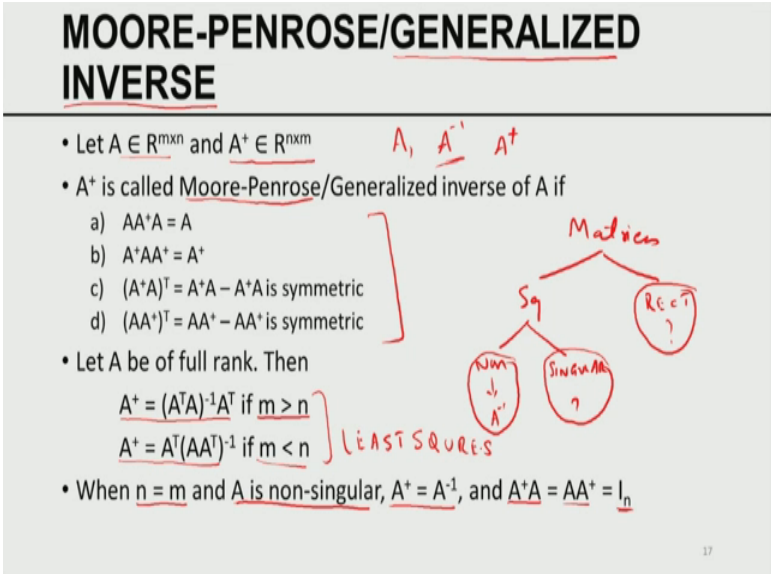
All the derivation are given. So, now, you can see I have derived the formula using these 3 pages of derivations.

(Refer Slide Time: 30:33)



So, this is one of the fundamental ways in which we can express the Sherman Morrison Woodbury formula. So, this proves the formula D in slide twelve by replacing B inverse by minus B by letting B is equal to 1, I can get the formula B C and A, now you can look at this. Now by this single one; single by proving this one single formula, I am able to look at this now by proving by proving D using this, this and this I come here then I specialized by choosing several different parameters I derive C, I derive B, I derive A. So, D is a generalization of A, B and C in that particular case, I am not going to go through the proof the proof is extremely simple in elementary and I would like you to try your hands on the proof.

(Refer Slide Time: 31:48)



So, that gives you the general aspect of what is called Sherman Morrison Woodbury formula the next one is the notion of generalized inverse, we already talked about inverses of matrices which are non singular mathematicians have been indulging in the concept of hey; how do I define inverses of matrices which are not square. So, let us talk about the basic ideas here.

Now I would like you to be able to. So, matrices square matrix rectangular matrices square matrix can be a singular non singular in the case of non singular matrix, I can define A inverse for singular matrix, I cannot define it for rectangular matrices, also we cannot apply this definition of a non singular matrix.

So, mathematicians have always been challenging themselves; how do I extend the notion of inverse of a non singular matrix to a general case of rectangular matrix to a general case of singular matrix and that is what is called generalized inverse generalized inverse is the generalization of the concept of inverse to matrices that are not necessarily square matrix.

The notion of a generalized matrix is very fundamental. So, if A is a matrix of size n by m by n; A plus denotes the generalized inverse of AA inverse denotes the ordinary inverse. So, I would like you to look at the symbolism AA inverse and A plus; this is the ordinary inverse this is generalized A plus is the generalized inverse Moore Penrose were the first one to define the notion of a generalized inverse. They said any matrix A plus that satisfies the properties A B C and D with respect to A is called the generalized inverse. In other words, AA plus A must be AA plus AA plus must be A plus A plus A transposes A plus A; that means, A plus A asymmetric AA plus transpose must be equal to AA plus A plus is must be symmetric.

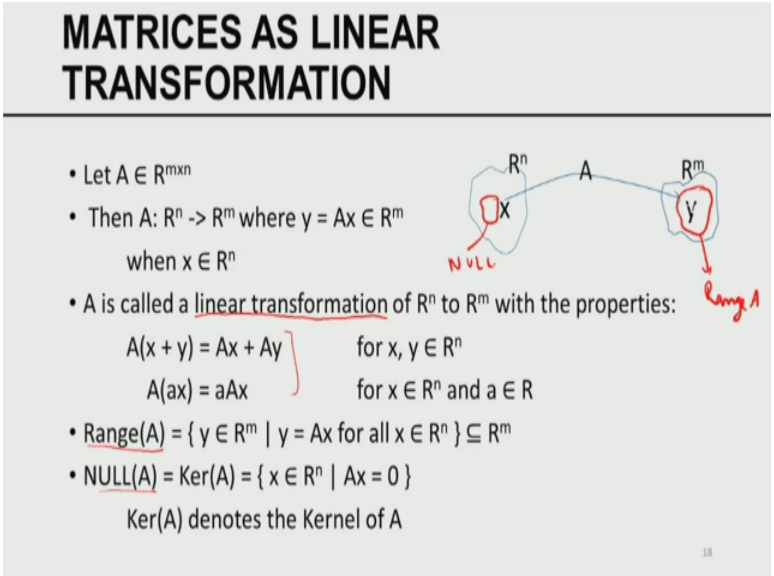
So, any matrix that satisfies these 4 properties can be regarded as the generalized inverse of A, it turns out if m is equal to n, A is non singular all these reduces to the definition of A inverse we already k. Now, therefore, this notion of a generalized inverse includes the ordinary inverse as a special case as a special case; in the case of rectangular matrices when a is a full rank what does it mean it is equal the rank of a is equal to the minimum of m and n in that case we can give specific formulas for the generalized inverse.

So, when m is greater than n the rank of a is said to be full rank if it is of rank n in that case A inverse is given by generalized inverse of a is given by this in the other case when m is less than n generalize inverse is given by this; we will talk about the occurrence and properties of generalized inverse when we talk about inverse problem in module 3.

Then I have already mentioned this, but it is worth repeating when m is equal to n and A is non singular non singular a plus b becomes A inverse in that case A plus AA plus becomes the identity matrix. So, this is a very beautiful mathematically consistent way of extending the notion of extending the notion of inverse of non singular matrices to rectangular matrices, it can also be extended to singular matrices in much similar fashion, but for cases the singular matrices, we do not have explicit formulas, but rectangular matrices with full rank, I have very specific formulas for generalized inverse again these are the basis by using which we will deal with lease squares theory the these generalized inverses occur very naturally in the theory of least squares sorry the theory of least squares.

So, these occur in the theory of least squares and we have seen in the morning lecture that yesterday’s lecture Gauss invented the LI theory of least squares; Gauss did not know at that time the notion of generalized inverse, but in 1930s; they had introduced this notion of generalized inverse and it turns out that generalized inverse and least squares theory least squares theory are intimately associated with each other. So, it is absolutely necessary that we have a nodding understanding of the Moore Penrose inversions properties.

(Refer Slide Time: 37:09)



Thus for we have seen several properties of operations on matrices special matrices. Now matrix can be also thought of as linear transformations of one vector space to another vector space. So, let A be a matrix of size m by n a, then as an operator maps the space R n to R m where y is equal to a times x here is the map here is an illustration this is R n, this is R m; A is the matrix which is m by n. So, if you take a n vector on multiplied by m by n matrix; I get a m vector. So, it transforms n vectors to m vectors and that transformation is induced by the matrix A. So, we call a an operator or a transformation the word operator and transformation are used synonymously.

We call an operator to be or a transformation to be a linear transformation or a linear operator from R n to R m, if it satisfies 2 properties A times x plus y is A x plus A y; A times A of x is A times A of x; if it satisfies these 2 properties the first property is called additively second property is called homogeneity. These 2 properties if a given matrix satisfies, then it is called a linear transformation.

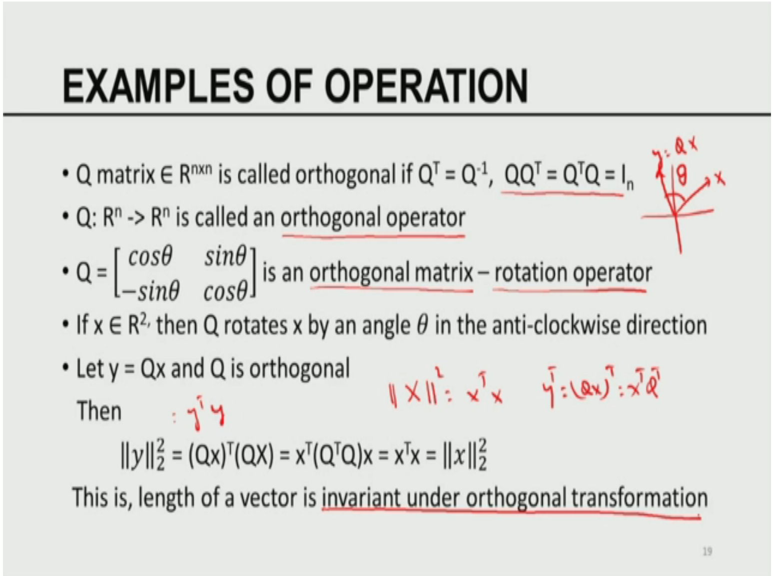
So, transformation linear transformation if there is a linear transformation; there should also be a non-linear transformation. So, transformations in general are of 2 types linear non-linear ee; it is a general property every linear transformation can be represented by a matrix that is a theorem in operator theory I am not going to going to that, but it is good to know.

So, given a transformation a there are 2 subspaces; there are 2 spaces associated with it one is called the range space another is called the null space there. So, given the range space consists of all those vectors y in R m where each y is obtained as a product of a and x for all x belonging to R of m. So, looking at this picture A is known; I pick x every one of them in R n and then I take every vector through A to this vector.

So, set of all collections y. So, obtained is called the range of A the null space of A; on the other hand is also called a Karnal of a these are different names is the set of all vectors which are annihilated by A. So, A times x 0, x belongs to R n is a set of A x is equal to 0 that is called the null space; it is also called a Karnal. So, the Karnal of A is the set of all vectors which are annihilated by the matrix A.

So, I would like to be able to emphasize that given a linear transformation a there are essentially 2 subspaces associated with it one is called the range space another is called the null space. So, the range space is a subspace of R n; the null space is a subspace of R n. So, if I work to talk about the null space is a subspace of R n range space is a subspace of R m. So, this is the range of A; this is the null space of a null space of A. So, you can see I am associating 2 spaces with a every given linear transformation.

(Refer Slide Time: 41:35)



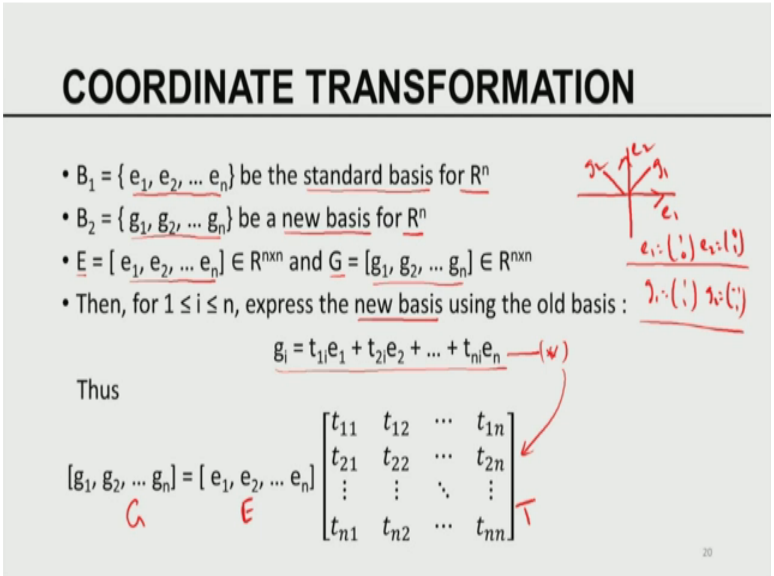
Now, I am going to talk about examples of a certain operations let Q be a matrix the matrix a is called orthogonal if Q transpose is equal to Q inverse Q maps R into R n is called an orthogonal operator where is a linear transformation as an operation op up operator; sorry I want to go back yeah is an orthogonal operator. So, Q is given by cosine theta sine theta sine minus sine theta cosine theta is a simple example of an orthogonal operator. These matrices also is also called orthogonal matrix orthogonal operator represented by an orthogonal matrix this matrix represents a rotation. So, what does it mean if you have a vector x if you multiply the vectors by Q.

So, if this is x if you have the vector y; y is equal to Q times x the length of x and length of Q are the same. So, this is called a rotation operator rotation operators are generally denoted by orthogonal matrices are orthogonal matrices represent rotation operators. So, if you multiply a vector in R 2 by Q you rotate the vector by an angle theta the theta is called the theta by which you rotate is called the cos is the theta that comes in Q cosine theta sine theta minus sine theta cosine theta are the 4 elements of the 2 by 2 matrix.

So, let y be equal to Q of x Q is called an orthogonal matrix then the norm of the square of y you already know the norm of x square of that is x transpose x likewise if you have square of the norm of y this is going to be equal to y transpose y, but y is equal to Q of x therefore, y transpose is equal to Q of x transpose the transpose of the product is the product the transposes taken the reverse order this is x transpose Q transpose.

Therefore, if I took the square of the norm of the vector y which is Q x transpose times Q x this is x transpose Q transpose Q of x, but by property of the orthogonal matrix Q; Q transpose is Q transpose Q is I; therefore, x transpose x that is the norm of x itself. So, if y is equal to Q of x and Q is orthogonal even though y is different from x they share the same length. So, orthogonal transformation preserves the length of the vector x. So, the length of the vector x is invariant under the orthogonal transformation and that is a fundamental property of the orthogonal matrices.

(Refer Slide Time: 44:54)



Now, I am going to quickly refer to coordinate transformations again the property of matrices linear transformation. Let us consider R of n; R of n is a basic R of n; sorry, yes, R of n is a space every space has a basis the standard basis for R of n is e 1, e 2, e n, these are called unit vectors given a space there are multiple basis for a given space for example, if you have R 2 e 1 e 2 is one's basis, I can also consider this is e 1 this is e 2.

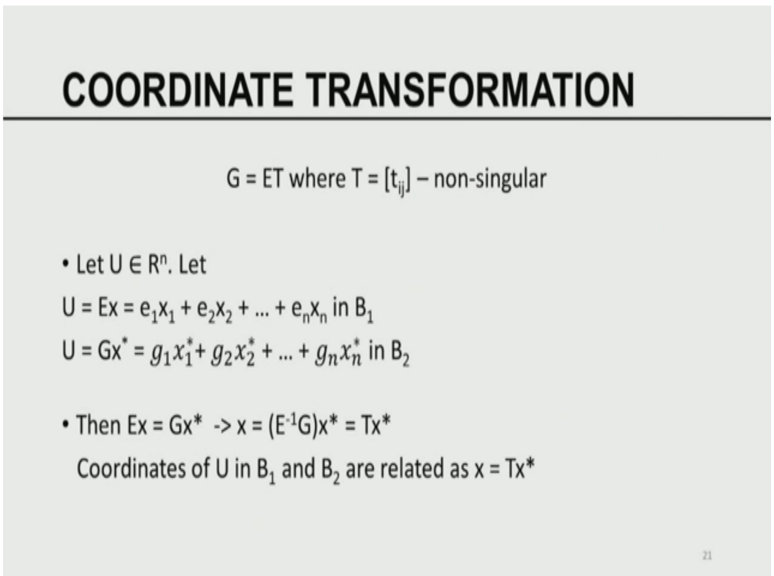
So, e 1 is equal to 1, 0, e 2 is equal to 0 1, I could consider this is this is g 1 this is g 2 what is g 1, g 1 is 1; what is g 2; g 2 is equal to minus 1; one that is the basis; this is the basis. So, given a space I can have multiple basis each basis has the same set of elements of vectors. So, if I am considering a space R n; I can consider a standard basis I can consider a new basis the standard basis are listed as e 1 to e n; the new basis are listed as g 1 to g n.

If I have a set of n vectors, I can create a matrix e e 1 to e n; if I have a set of vectors g i can construct a matrix g, g 1 to n. So, e is a matrix consisting of standard basis vectors g is a matrix consists of the new basis vectors both of them are basis both of them span a or the same equal to R of m. So, it behoves to ask a question how do these 2 basis e and g are related to each other we are going to show that these 2 bases are related by a linear transformation and how do you show that every element to the new basis g can be expresses the linear combination of the elements on the old basis because every vector the new basis is a vector n R n; R n has unit basis the standard basis. So, I can express g as the linear combination of the standard basis.

If I did this for every g, g 1, g 2, g n, this is the general expression for g I; I can now collate all of them. So, if I consider g 1, g 2, g n, in the form of a matrix please understand each g is a vector. So, first vector second vector nth vector; this is the matrix this is the matrix g, this is the matrix e, the matrix e and g are related through the elements ts; it can be easily verified that this relation is very fundamental this relation induces simultaneously.

So, this is for one g i; if I consider all the g i's together this is the resulting relation. So, now, you can see g is related to e; g is related to e; through this matrix t. So, you can think of this as t is a transformation that relates the basis g 1 that relates the new basis with the old basis.

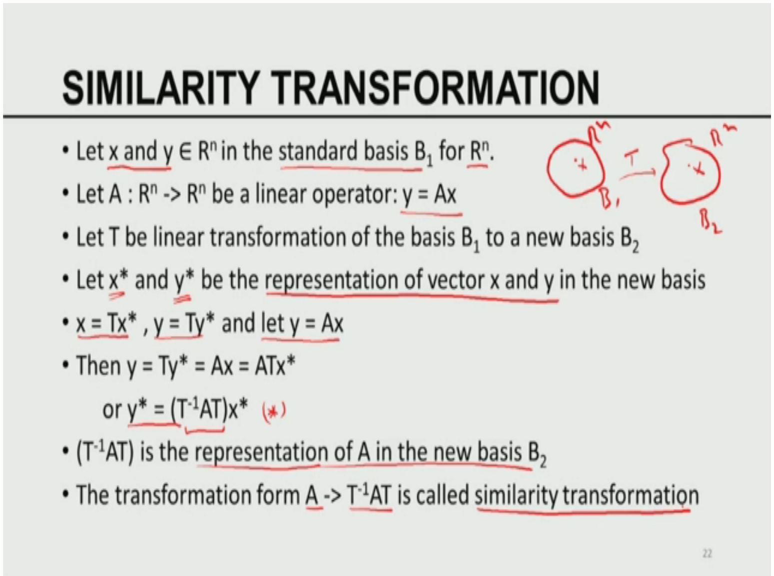
(Refer Slide Time: 49:04)



So, we can yeah; we can denoted as the new basis equal to the old basis times. So, the new basis equal to the old basis times T and t is the transformation and it can be shown that transformation is not singular.

So, what is the role of a linear transformation role of one of the roles of linear transformation is that it transfers one set of basis of a given vector space to another set of basis and they are related linearly through the linear transformation T to the linear transformation T. So, this essential tells you the coordinates of the new basis and the coordinates of the old basis are related and.

(Refer Slide Time: 49:55)



And that brings us to a special class of transformation called the similarity transformation. So, I am now going to be talking about space R n; I am going to be concerned with the standard basis b 1 instead of calling it e; I am simply going to call it B 1. So, let x and y in R and B the standard basis elements in the standard basis R of n, A is a map from R n to R n we can think of this as the linear operator y is equal to A of x.

So, now let us think of it; now I have R n in 1 basis; I have R n in another basis; I have R n in another basis. So, the basis for this is B 1; the basis for this is B 2 and we know that there is a transformation linear transformation t that can map the basis B 1 to B 2, we have already seen B 1 was e, B 2 was g; remember that that been the last slide. So, T is the linear transformation from basis B 1 to the new basis B 2 let x star. So, if I consider a. So, both R ns are the same space even though I have drawn it differently if there is a vector x, here I have the same vector x here this vector has a representation in B 1 this representation this x has a representation in B 2; if the basis are related, I can also relate the representations of x in these 2 basis.

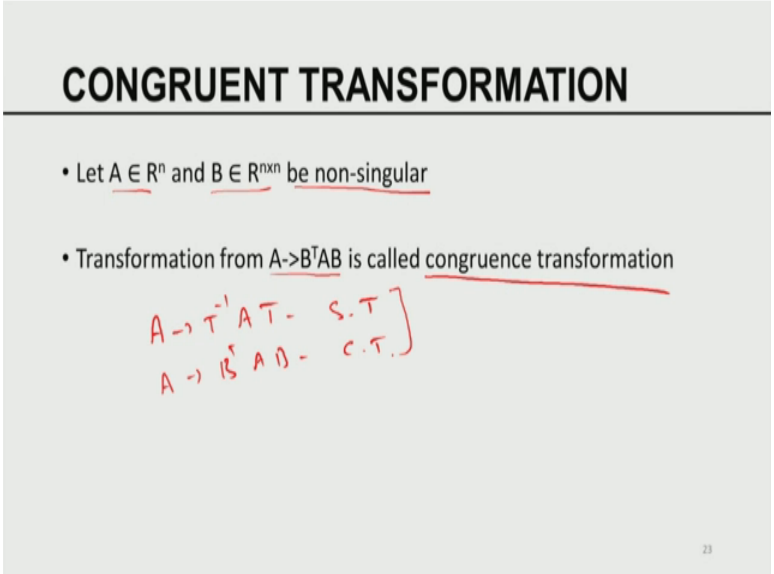
So, let x star be a representation of the vector x in B 1 let y star B, the representation of the vector y in the new basis. So, if I have Ts x is equal to T times T of x star y is equal to T of y star; let y is equal to a of x; I hope you understand that lots of animals in here.

So, x and y are 2 vectors in R n the standard basis A is a linear operator that transforms y to x t is the linear transformation from the basis B 1 to the new basis B 2. So, I have a linear operator; I have a basic transformation; I have then the representations of vectors from x to y the new basis, if I put them all together x is equal to T of x star y is equal to T of y star; let y is equal to a of x that essentially tells you y is equal to t of y star which is equal to a of x, but x is equal to T of x star, but y is equal to t of y star. So, if I substitute y is equal to t of y star we get the relation y star is equal to t inverse AT x star. So, you can readily see the relation between y; y star and x star as follows.

So, we have not changed anything we are simply concerned with 2 vectors x and y we are simply representing x and y in 2 different basis of R n and. So, A is an operator, T is the transformation; all these things relate to the fund reach to the fundamental result which is given by the star. So, I have a new matrix the new matrix is t inverse AT it is related to AT inverse AT is the representation of a in the new basis B 2.

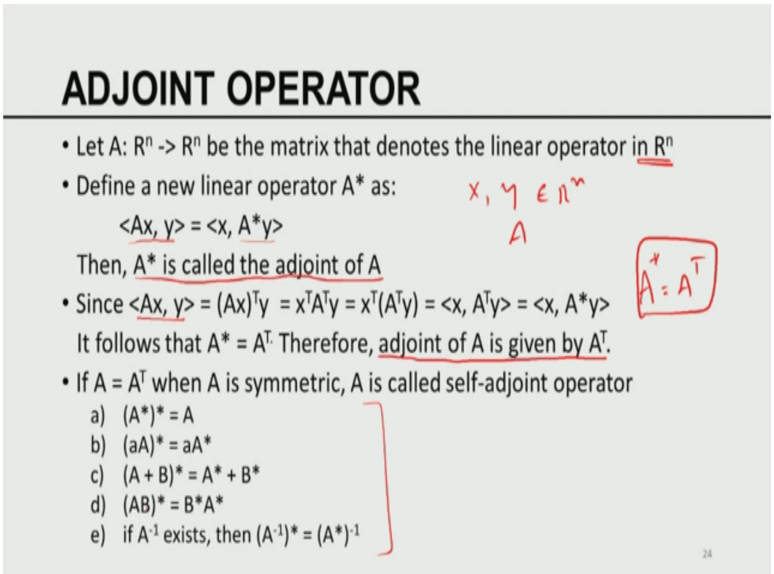
So, this transformation of the matrix A to T inverse; AT is called the similarity transformation it again plays a fundamental role in linear algebra. So, similarity transformation is a special class of linear transformation when you represent 2 vectors in a given vector space and these 2 vectors are related by an operator a if I change the basis for the same space from B 1 to B 2; then there is a transformation vector t comes into play. So, t a together help us to be able to define the linear transformation. So, this repra; this is called representation of matrices in different basis a is a representation in one basis t inverse AT is a representation of the same matrix or same operator in another basis; these 2 are related by the fact that care calls the at a transformations.

(Refer Slide Time: 54:47)



The next transformation is called congruent transformation let A be a matrix, B be a matrix, I am requiring B to be non singular, A is any matrix a transformation from A to B transpose; AB is called a congruence transformation, please remember the differences. Now A to T inverse AT that is similarity transformation A to B transpose AB that is called the congruence transformation congruence transformation. So, these are 2 transformations of matrices that occur very naturally in linear algebra. The reason we are talking about congruent transformation and similarity transformation because these are special cases of linear transformation.

(Refer Slide Time: 55:46)



Now, I am going to define at other concept called the adjoint operator; any of you who have done 4 D-VAR; you will know 4 D-VAR is called also called adjoint method adjoint is a property of operators that comes essentially from matrix theory operator theory. So, I am going to quickly define the properties of adjoint operators which are fundamental to understanding data assimilation method called 4 D-VAR.

Let a be a matrix that denotes the linear operator R n. So, a matrix; matrix can be called as an entity as a matrix and define the operations on that a also a matrix can also be thought of as an operator effect on vectors on R n. So, the same object plays 2 different roles either a matrix or as an operator a matrix is representation of an operator now define a new linear operator a star and the definition goes like this given 2 vectors. So, let us talk about this. Now I am having R n to start with; I am having an operator A and R n.

So, what does a and R n means it takes vectors in R n and maps to vectors in R n. So, let us speak 2 vectors x and y belonging to R n. I have been given a matrix A, if I have x and y and a, I can compute the matrix vector product a x that is a vector I, y is a vector, I can compute the inner product. So, this is the inner product of a x and y the inner product of a x and y is related to inner product of x with a star y. So, what does it mean? A x is a transformation of A star y is a transformation of y the matrix A star that forces this equality is called the adjoint of A; it is called the adjoint of A. So, that is the definition of the property adjoint.

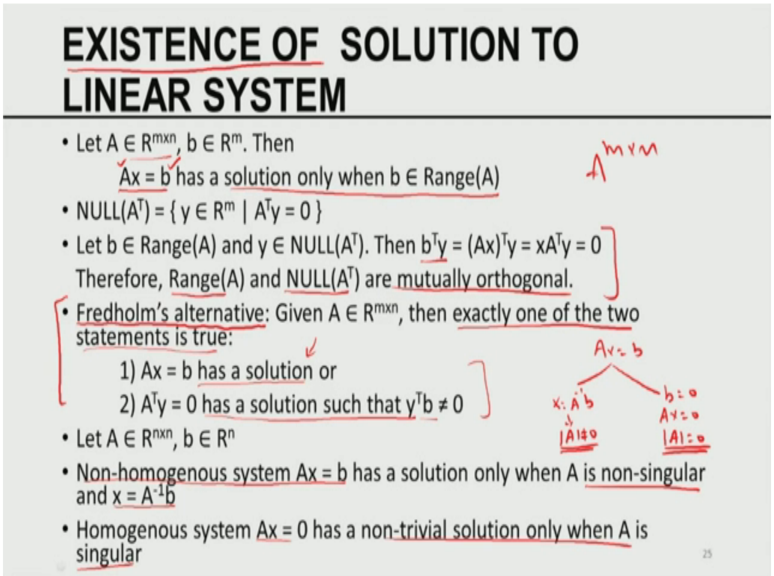
This adjoint a is not unknown to us if you look at the standard definition inner product if you consider a inner product of a x and y by definition inner product of a x and y is a x transpose y, but a x transpose y is x transpose A transpose y x transpose A transpose y is x transpose times A transpose y; I can associate like this; this can be express as the inner product of 2 vectors and that can be expressed as x is equal to a star y therefore, in general the adjoint of a matrix is the transpose the transpose is the adjoint. So, that is the fundamental thing that comes from this analysis.

So, for finite dimensional vector spaces if you are considering matrices of finite dimensions the transpose operation is related to adjoint. So, transpose is a urinary operation, we have already defined a simple operation adjoint is a another concept it turns out adjoint can be represented as transposes in this particular case of matrices, but adjoint in general is a much more deeper fundamental concept in operator theory in operator theory.

So, adjoint of a matrix transpose of a matrix these are urinary operations on operators are matrices adjoint of an adjoint is the original matrix. So, adjoint of an adjoint is AA times A, adjoint is AA times adjoint of A adjoint of a sum is the sum of the adjoints; adjoint of a products is the product of the adjoint taken in the reverse order. If A inverse exist adjoint of A inverse is the inverse of the adjoint; these are very fundamental properties of adjoint with respect to other operations.

So, how adjoint behaves with respect to adjoint how adjoint behaves to the 2 scalar multiplication how adjoint behaves with respect to matrix addition matrix product and inverse. So, the interaction of 2 different operations is the topic of discussion in here the notion of an adjoint operator and its close relation to transpose.

(Refer Slide Time: 60:21)



Now, we come to one of the fundamental concepts in linear algebra why do we do all these things ultimately we would like to be able to solve the equations. So, given an equation A x is equal to b under what condition A x is equal to b as a solution. So, we are interested in analyzing the existence of solutions of linear systems let A be a m by n matrix when m is equal to n a special case.

So, we are going to start with general matrices. So, A x is equal to b, A is known, b is known, I want to solve the inverse problem; I want to find an x before you compute an x, you have to verify the solution you have to assure yourself the solution exists. In this case, A x is equal to b has a solution only when b lies in the range of a you remember range of a we have already defined range of a is a set of all vectors that a maps from the domain to the co-domain.

We also have talked about null space of A, if we can talk about null space of A, I can talk about the null space of A transpose. So, null space of A transpose, a set of all y belonging to R n such that A transpose y is 0 is 0. So, if b belongs to the range of A and y belongs to the null of A, then b transpose y is A x transpose y is x transpose A y is equal to 0 either is a property that follows from the fact that y belongs to the null space of A.

So, what does this tell you x belong b belongs to b belongs to the range of A y belongs to null of a this is the inner product of 2 vectors one from the range another from the null space the inner product of 2 vectors 0 means orthogonality. So, this essentially tells you the range of a and the null of a are mutually orthogonal.

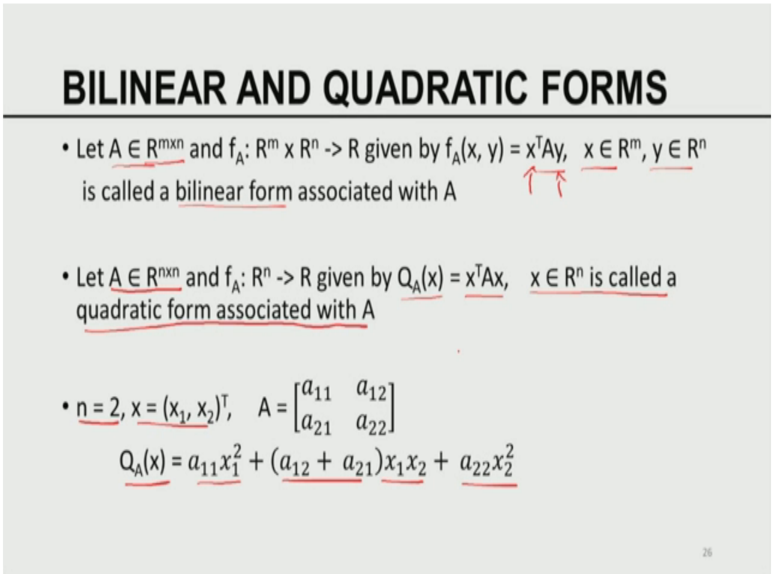
Please remember this is a fundamental property given a matrix A of size m by n, we have associated 2 spaces the range space the null space this essentially tells you the intrinsic property of the behaviour of vectors one from the null space another from the range space they are mutually orthogonal now coming back to the existence question there is a famous result by Fredholm Alter is called Fredholm's alternative Fredholm's alternate; essentially says the following given a matrix A which is m by n, then exactly 1 of the 2 statement is true; either A x is equal to b as a solution or A transpose y equal to 0 has a solution such that b transpose y are y transpose b is not equal to 0.

So, these are the only 2 possibilities that can happen for a general case of matrices which are rectangular when m is equal to n; A belongs to R n by n b belongs to R n then the non homogeneous system of equation A x is equal to b has a solution only when a is non singular and x is A inverse b that again follows from the alternative A of the Fredholm alternative; the homogeneous system A x is equal to 0 has a non trivial solution only way is singular. So, these are the 2 basic fundamental facts. So, what does this say? If I want to be able to solve A x is equal to b, I have a unique solution x is equal to A inverse b when a is non singular the determinant of A is not equal to 0, in this A b or else what happens the homogeneous system in this case b is equal to 0 x is a non trivial case.

In this case, A x is equal to 0 has a non trivial solution only when a singular. So, in this case the determinant of a is singular; these are the 2 fundamental differences. So, a homogeneous system has a non trivial solution, when the matrix a is singular, a non homogeneous system has an non trivial solution when the matrix is non singular and these 2 are consequences of the fundamental property call Fredholm's alternative and these 2 together provides condition for the existence of solutions of linear system.

I am not going to prove the uniqueness you can always if the matrix is non singular; A x is equal to b not only the solution exist we can also show the solution is unique once you know the solution exist and is unique we can then try to develop computations to be able to actually develop the solution to show something exist is one thing to be able to derive or build or computed is something else, but to be able to compute; I must have been assured that the solution exist.

(Refer Slide Time: 65:49)

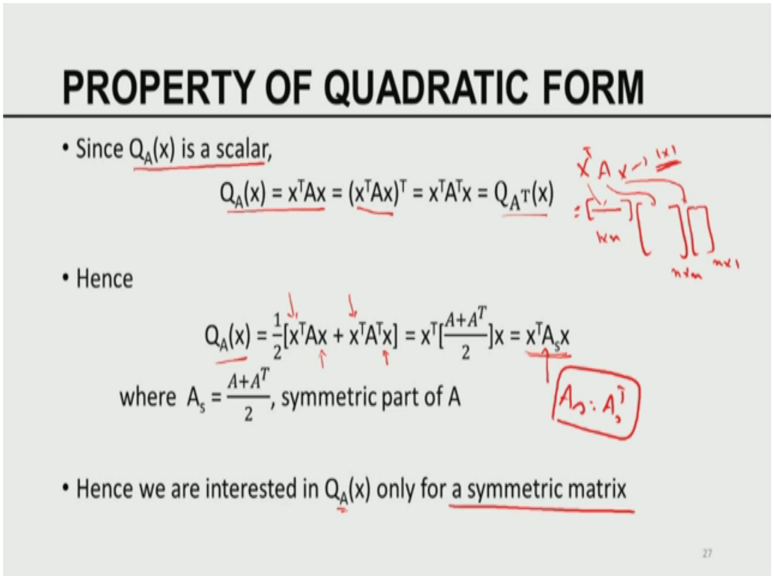


The next set of ideas from matrix theory are called bilinear and quadratic forms let A be a matrix of size m by n; sorry, let A be a matrix of size m by n, I am given 2 vectors x is in R m y in R m, I can define a functional defined by a which is f; f A x of y x transpose A x that is called a bilinear form; bilinear is a linear in x, it is also a linear in y. So, because it is linear in 2 variables at a given time it is called bilinear.

When as, but when a becomes instead of rectangular matrix by a square matrix n by n; I can define QA of A as A transpose A x for x in R n and this is called a quadratic form associated with the A. So, this is a bilinear form is a first degree in x and y here quadratic form is of second degree in x bilinear forms are linear in each of the variable quadratic forms are quadratic in the components of x.

So, here is an example of a quadratic form let n be equal to 2 x be a vector x 1, x 2, let A be a matrix given by this, QA of X is equal to a 1 1 x 1 square plus a 1 2 plus a 2 1 x 1 x 2 plus a 2 2 x 2 square you can see the first term is quadratic in x 1 this is the quadratic in the product x 1, x 2, this is the quadratic in x twos. So, this is an example of what is called a quadratic form bilinear form quadratic forms are special cases of bilinear forms.

(Refer Slide Time: 67:52)



What is the property of quadratic form? Let QA be a scalar in this particular case, we already know we already know QA is a quadratic form, QA is given by X transpose A x. QA is a scalar X transpose A x means what X transpose A x x is A x is a row vector, A is a matrix, then I have a column vector. So, x is a column vector, A is a matrix; this is a row vector. So, this is 1 by n; this is n by n; this is n by 1. So, the whole thing is 1 by 1; one by one is a scalar you all know that basically a quadratic function is a scalar the transpose of a scalar is itself.

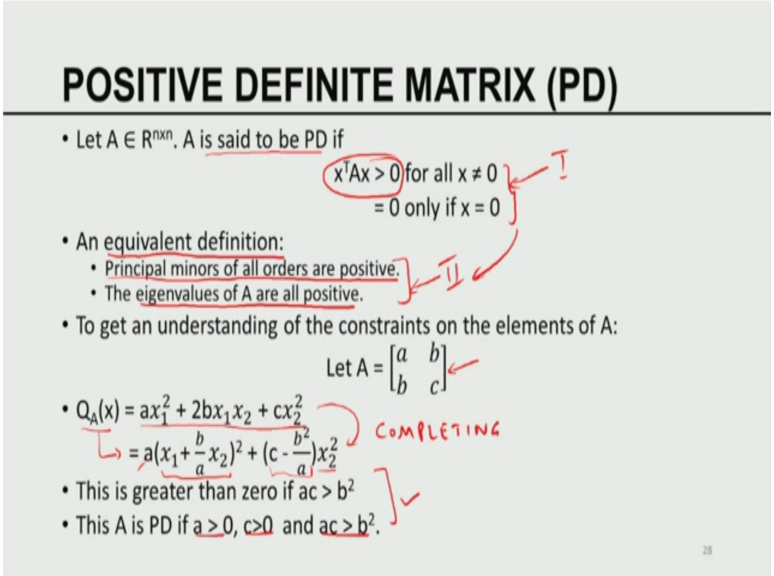
So, I can the scalar is its own transpose, but transverse is the product is the product of the transpose has taken in the reverse order we have already seen the product of the urinary operation transpose. So, this is equal to x transpose A transpose x. So, this is the quadratic function in Q of A transpose f x. So, what is that we have shown a quadratic form of A is same as quality form of A transpose that is a fundamental property.

So, if these 2 are equal, I can then write QA is equal to 1 half of the sum of QA and Q A transpose because these 2 are equal. So, this is equal to one half of the sum of this and that this is equal to x transpose a this is equal to x transpose A transpose A, I can do a little bit of an algebra in here x transpose is the left common variable x is the right common variable I can take the right common left common I can arrange the inner matrix as a plus a transverse by 2 you will quickly recall, then we talked about decomposition of matrix and symmetric and skew symmetric part this is the symmetric part. So, Q of A of x is the same as x transpose A of s times x. So, this is called the quadratic form related to the symmetric part of x.

So, if you are interested in quadratic forms we can without loss of generality assume the matrix A is always symmetric, if it is not, I can convert the matrix A to its symmetric part, I have not changed anything because symmetric part of matrix is always symmetric a s is equal to a s transpose and this property is routinely used in data assimilation techniques, again these are all fundamental properties that come from matrix theory.

So, quadratic form the quadratic form with respect to a vector each term consists of second degree term as we saw in the previous example x 1 square x 1, x 2, x 2 square; this quadratic form has a special property the special property being the fact that quadratic form of A is the same as quadratic form of the symmetric part of A because of this from no one without loss of generality when we are going to assuming quadratic forms we will only take symmetric matrices.

(Refer Slide Time: 71:28)



Once you have the notion of a quadratic form, I can then find split the idea of quadratic forms one is called positive definite quadratic form. So, this is a fundamental concept; again, let A be a matrix, A is said to be positive definite is a definition; if x transpose A x is greater than 0 for all x not equal to 0 is equal to 0 only when x is equal to 0. So, this is the definition of a quadratic form. So, what does it mean in general the quadratic function of a matrix need not be positive, but if A is positive definite in the quadratic function always is positive?

So, this is the further definition of positive definite functions or pos positive definite quadratic forms I hope that definition is very clear. So, I want do to go back. So, we simply define the quadratic form to be the we define the quadric forms to be given by this product in here there is no condition on the sign of this except this is a scalar now what is that we are saying this one is not only a scalar, but also it takes a positive value for all x; x can have positive negative elements he also can have positive negative elements, but if you consider the product x transpose a x its always positive when x is not equal to 0; it is 0 only if when x is equal to 0.

This positive definite quadratic forms have different ways of can be explained in different ways; there are equivalent definitions for quadratic forms one definition of quadratic form is what we defined, but this definition is not very useful to apply; if somebody gives a verdict form, if I want to apply this definition; I have to test it for infinitely many xs it is not possible. So, this is a very nice definition, but computation will not useful.

So, there are equivalent definitions which are computationally meaningful. So, tests have been developed to decide under what condition a matrix is positive definite one of the conditions is if the Eigenvalues of a are all positive then it is positive definite if all the principal minors; if principal minors of all orders of a are positive the matrix is positive definite a principal minor is the determinant of a sub matrix a matrix has several different principals of minors. So, if all the principal minors of a given matrix are all positive that been the determinants of all possible sub matrices in a given matrix are positive they matrix is positive definite.

So, these 2 definitions give you an algorithmic way to test for positive definiteness this is simply a fundamental definition this first definition is not very useful in terms of computation the second view is derived from the first view, but it is very useful computationally to get an understanding of the constraint on the elements of A to be a positive definite matrix. Let us consider a symmetric matrix a b, b c please understand with respect to with respect to matrices in the context of positive definiteness we need to consider only symmetric matrices. So, we can consider symmetric matrices like this. So, if you consider QA of x for this matrix; this is this takes this form; I can rewrite this by simply completing the squares like this a simple algebra I will show you like this; this is called the method of completing the square.

So, by method of completing the squares we can express the expression for a quadratic form like this; now I would like to examine this expression what are the condition necessary in order to make this positive as required in the in the in the condition one a square of any number is always positive. So, this term is always positive. So, in order that this term is positive I have a positive x 2 square is always positive. So, in order that this term is positive; I have to ensure that this term is positive.

So, we can state that we can state that a is positive definite in this case, if a is positive is c is positive if a c is greater than b square if a c is greater than b square this is positive a is positive this is positive, I could have rewritten this by completing the square with the other way that will give you c is positive.

So, a matrix of this type is said to be positive definite if a is positive c is positive a c is greater than b square. So, this is an example with this condition of a positive definite matrix of a positive definite matrix. So, you can see not every matrix is positive definite positive definite brings constraints and the elements of the real values of the values of the matrices.