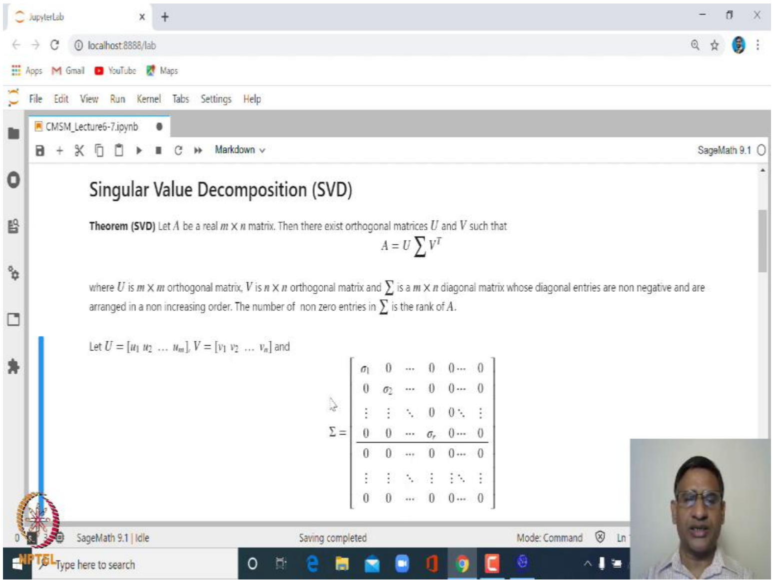
Computational Mathematics with SageMath
Prof. Ajit Kumar
Department of Mathematics
Institute of Chemical Technology, Mumbai

Singular Value Decomposition (SVD)
Lecture – 41
Singular Value Decomposition (SVD) with SageMath

(Refer Slide Time: 00:15)



Welcome to the 41st lecture on Computational Mathematics with SageMath. In this lecture, we will look at a very important concept known as Singular Value Decomposition of a matrix.

In short, we write this as SVD. This SVD has several applications in science and engineering.

So, let us start with, what do I mean by this Single Value Decomposition. Let me state the result. Suppose, you have a matrix A, which is a real m cross n matrix and, then it says that you can find orthogonal matrices U and V such that; A can be written as U into sigma times V transpose.

Here, U and V are orthogonal matrices and sigma is a rectangular matrix, but it is a diagonal matrix and these non-zero entries of sigma are arranged as decreasing order of the magnitude.

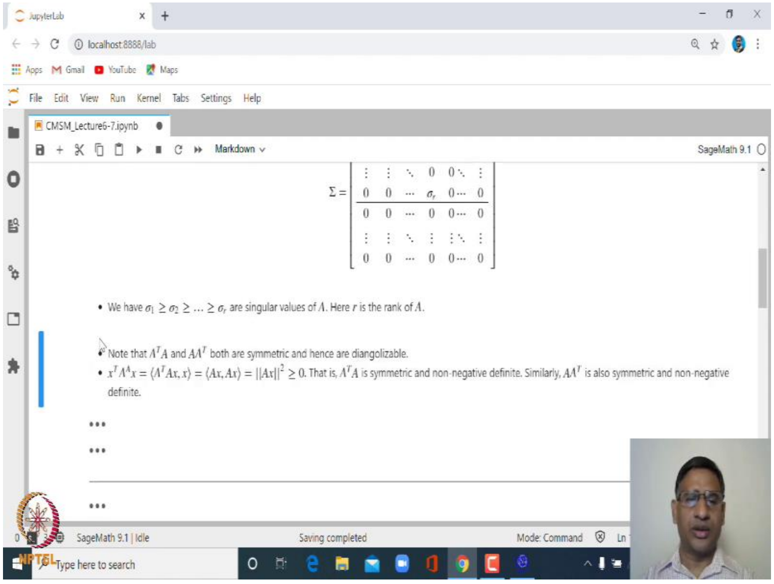
Let us let us see what do I mean by that. Here U, you can denote it by let us call a matrix u1, u2, um and since U is orthogonal, these u1, u2, um will be orthonormal set of vectors.

Similarly, V, we can denote it by v1, v2, vn , and sigma is given by sigma1, sigma2, sigmar and then all zeros. Here r is the rank of the matrix.

So, in case rank is r, the only non-zero entries will be sigma1 to sigmar and this sigma1 is bigger than equal to sigma2, sigma2 is bigger than equal to sigma3 and so on.

This is what you have as a non-zero diagonal matrix. The remaining all entries are 0. That is how this sigma, U and V look like. Here the sigma1, sigma2, sigmar are also known as singular values of A, and the number of singular values of A is nothing but the rank of this matrix A.

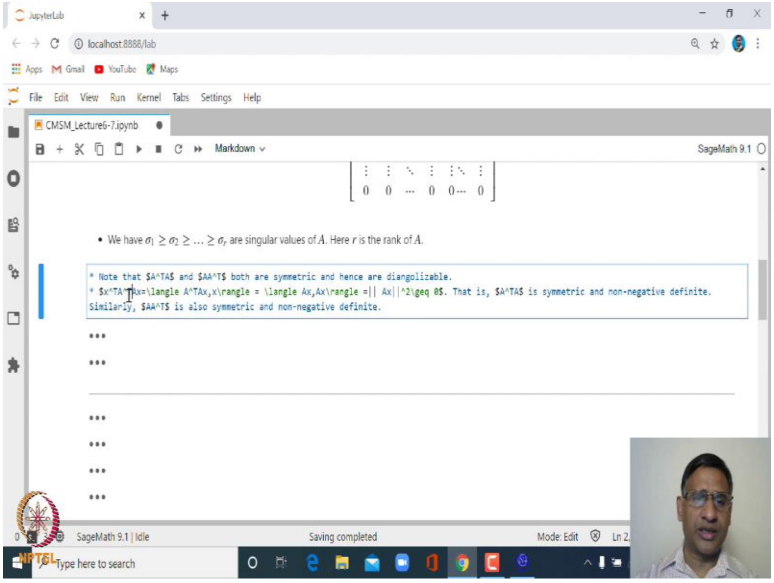
(Refer Slide Time: 02:23)



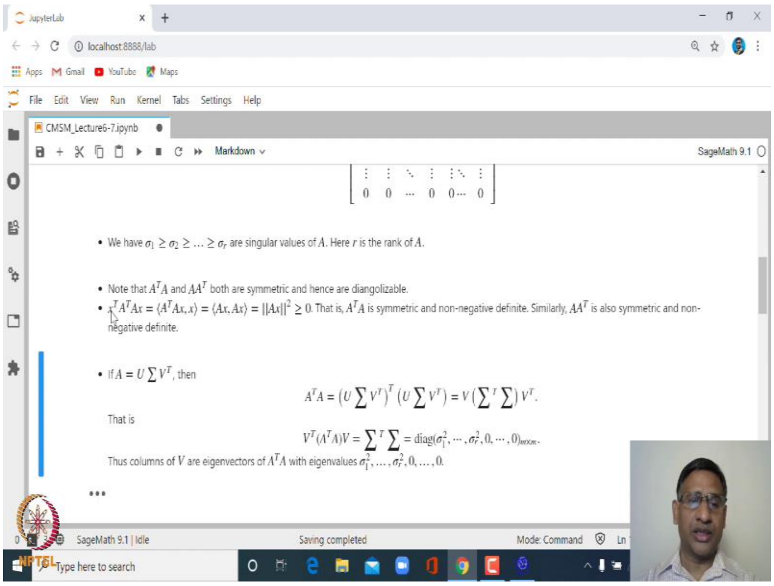
Now, let us play with this singular value decomposition. A is equal to U into sigma into V transpose. Suppose we look at what happens to A transpose A and A A transpose.

Notice that both A transpose A and A A transpose are symmetric matrices and therefore, it will be diagnosable. Not only that, if you look at for any vector x, x transpose A transpose A x.

(Refer Slide Time: 03:02)



(Refer Slide Time: 03:07)

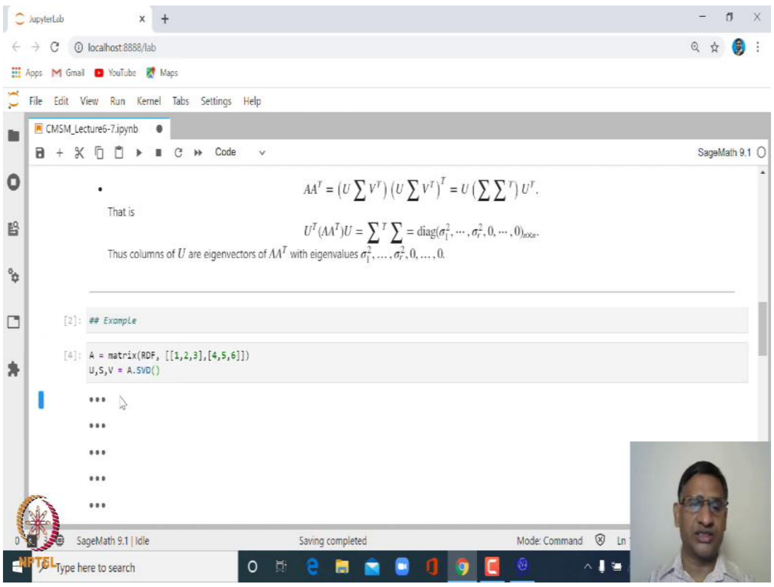


So, this would be A transpose A this is nothing but inner product of A transpose A x with x which is same as saying inner product Ax with Ax, that is nothing but norm of Ax square, which is always non-negative. So, for any vector x, what we have, x transpose A transpose Ax is always non-negative. That is same as saying, A transpose A is actually a symmetric and non-negative definite matrix. In particular, all eigenvalues of A transpose A will be non-negative. Similarly, you can check the other one. Let us say y transpose A into A transpose y, that also you can check.

Now, let us look at. Suppose A is equal to U into sigma into V transpose. Then if we compute A transpose A, then just replace the formula for A, what you get is this, This is nothing but V into sigma transpose sigma into V transpose. Now, what does it say? If you just multiply by V inverse on the left hand side and, then V on the right hand side, since V is orthogonal, V inverse is nothing but V transpose. So, what you have is, V transpose A transpose A into V is sigma transpose sigma. But sigma transpose sigma is actually nothing but a diagonal matrix sigma1 square, sigma2 square up to sigmar square and then, all the remaining 0's. This is m cross m diagonal matrix. That means, the A transpose A is diagonalized using V, that is same as saying the columns of V which are v1, v2, vm are nothing but eigenvectors of A transpose A, with eigenvalues sigma1 square, sigma2 square and so on.

Then, take the square root of the sigma1 square, sigma2 square, non-negative square root of these, that is how you obtain this matrix sigma.

(Refer Slide Time: 05:24)



Similarly you can look at what happens if you calculate A into A transpose. When you calculate A into A transpose, what you will get is; U into sigma sigma transpose U transpose. Therefore, you transpose A A transpose U is nothing but diagonal matrix sigma1 square, sigma2 square up to sigmar square and remaining all 0, this is n cross n diagonal matrix which has only r non-zero entries. That means that the columns of U are eigenvectors of A A transpose with eigenvalues sigma1 square, sigma2 square, up to sigmar square. Therefore, in order to find U and V, all we need to do is we need to find these eigenvectors of A A transpose and A transpose A.

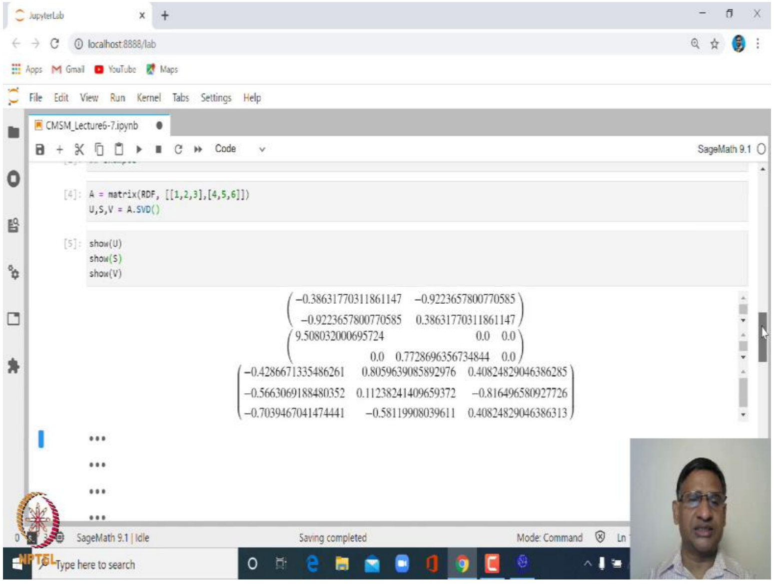
Eigenvectors of A transpose A are nothing, but columns of V. Eigenvectors of A A transpose are nothing but what column vectors of U. Of course, you need to find eigen basis of A A transpose and A transpose A and that eigen bases will be actually the columns of U and V. Also, since A transpose A and A A transpose are symmetric matrices, eigenvectors corresponding to distinct eigenvalues will all be orthogonal. That is why we are able to find U and V which are orthogonal matrices with this property. So, this is a very simple concept and actually the proof of this theorem is also very easy. This is actually a constructive proof.

You simply need to start with eigenvectors of A transpose A and then, accordingly you define vectors ui right. So, if you are interested to look at the proof of this result, you can look at any standard book on applied linear algebra.

Now, let us take an example. Suppose we have a matrix A, and since we wanted singular value decomposition that involves finding eigenvalues, eigenvectors, you may have to find numerical eigenvalues, eigenvectors. So, we will take the domain as RDF. So, the matrix is 1 2 3; 4 5 6. So, that is the matrix. This is a rectangular matrix and it has rank 1.

So, let us see. Sage has inbuilt method to find singular value decomposition. All you need to do is, A dot SVD an empty round bracket.

(Refer Slide Time: 08:25)

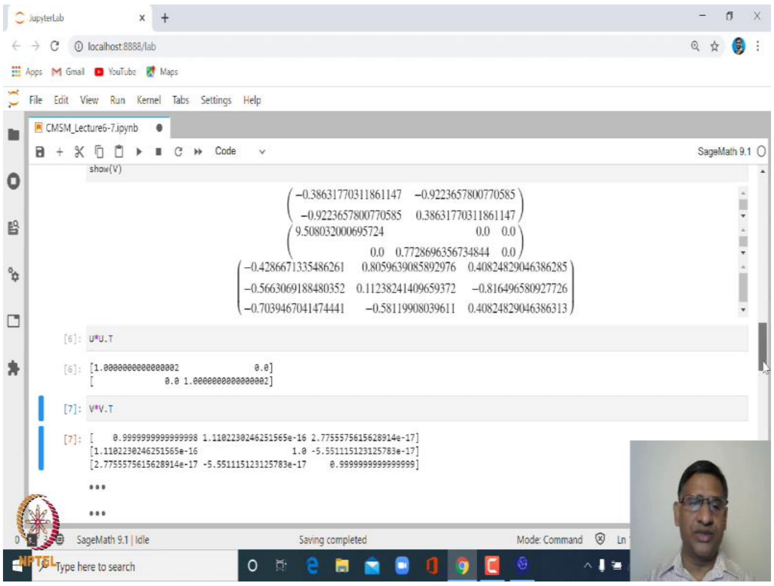


Let us compute this and let us ask it to show what are the U S and V, I am denoting it by S. This is the matrix U, which is a 3 cross 3 matrix namely. Notice that A is 3 by 2 matrix and therefore, and this is the sigma, is actually 2 cross 3 matrix.

And these are the diagonal entries, this is going to be actually the square root of the eigenvalues of A A transpose or A transpose A and this is the second one.

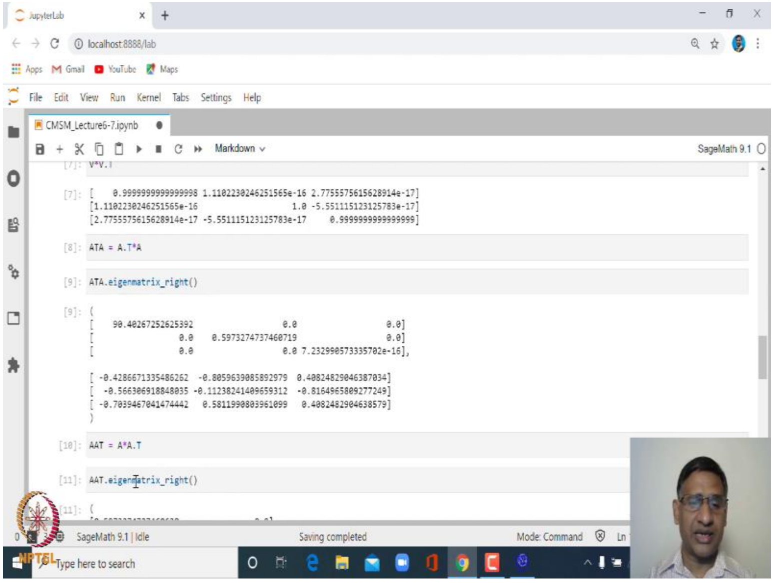
And also, you can see here this is arranged in decreasing order. This is the maximum eigenvalue, this is the smallest one and this is V. V will be 3 cross 3.

(Refer Slide Time: 09:20)



And you can check that, U and V, what we have obtained are orthogonal matrices. So, if I look at U into U transpose, that should give be identity matrix. Similarly, if I look at V into V transpose, that should also be identity matrix. Of course, you can see here, this entry is of the order 10 to the power minus 6 which is as good as 0. This is 10 to the power minus 7, which is as good as 0.

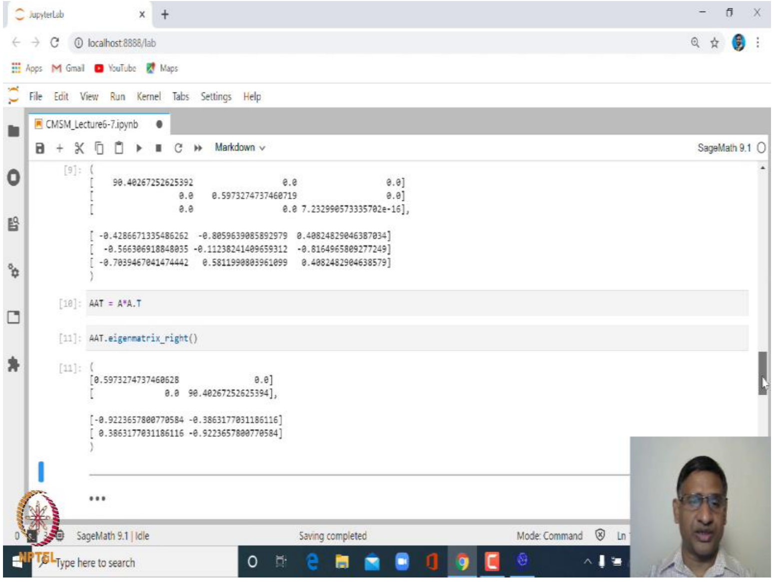
(Refer Slide Time: 09:50)



So, U and V are orthogonal matrices. Next, let us look at how we can obtain this manually. First, we need to compute A transpose A. Let me denote A transpose A as ATA and once we have defined this. Let us find eigenmatrix of A transpose A. So, eigenmatrix will give you two matrices.

One is the diagonal matrix, since this matrix is diagnosable, this will be the matrix D whose entries are eigenvalues. So, this is a sigma 1 square, this is sigma 2 square, this is sigma 3 square, this is 0. this is the matrix of eigenvectors. Similarly, you can find out A into A transpose and find out eigenmatrix of A into A transpose.

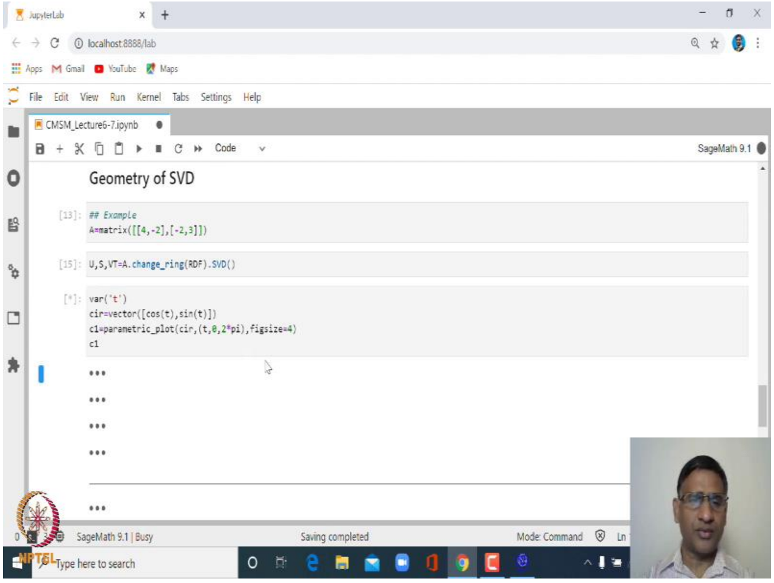
(Refer Slide Time: 10:50)



So, this is what you get. This is the diagonal entries, these are the eigenvalues. Of course, here it need not be arranged in decreasing order of eigenvalues, but this is what you get. Now, you have to interchange this to two columns.

You can just try to compare what we have obtained eigenmatrix in this case and what we obtained using the inbuilt function. You can see here, this this first column is same as first column of V, which we obtained using inbuilt function and the second column and third column will also be the same.

(Refer Slide Time: 11:35)



So, next let us look at what is geometric meaning of the singular value decomposition. You see A is a linear transformation. If A is m cross n matrix, A is a linear transformation from Rn to Rm. We would like to know how any object gets transformed by this matrix A.

For example, if I take any vector what happens to that vector. Since A is split as U into sigma into V transpose, A of any object will be U times sigma times V transpose. Now, U and V are actually orthogonal matrices and orthogonal matrices are nothing but actually matrix of rotation.

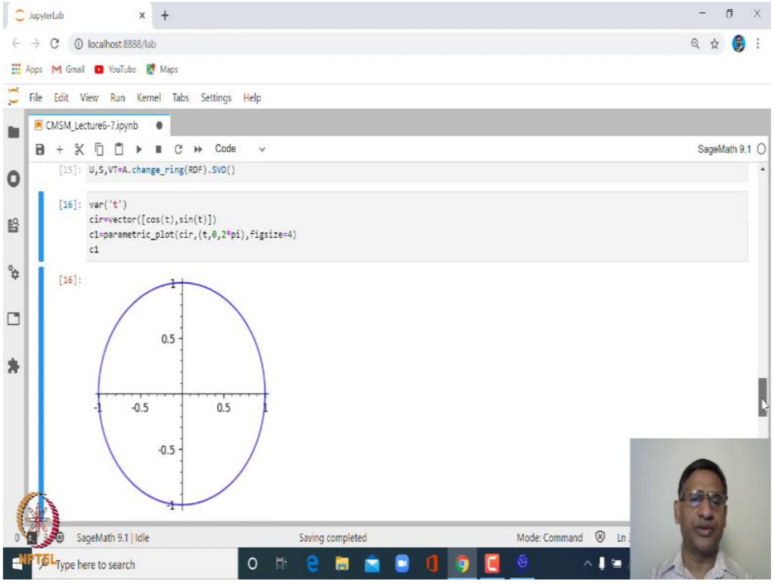
What V transpose will do? It will rotate that vector and sigma is a diagonal matrix basically. So, that is going do what? It is going to scale in each coordinate direction and then you rotate by U. So, any action of A onto any vector can be obtained as the sequence of action by V transpose sigma and U transpose.

That gives you actually geometry, how an object changes by a matrix A.

Let us try to demonstrate this in 2-Dimension and 3-Dimension. Suppose you have a matrix. Let us look at this. This example we have already seen. Let us go back to this place. We have a matrix A which is 2 by 2 matrix 4 minus 2 minus 2 3, this is a symmetric matrix.

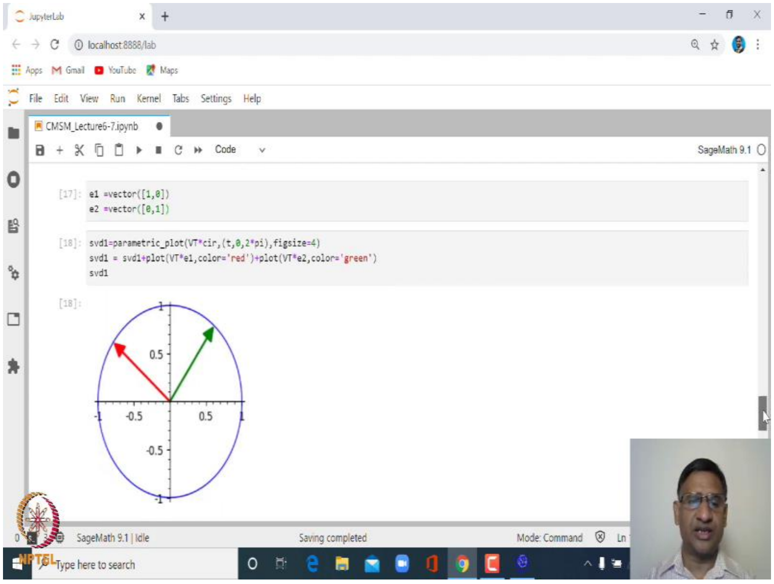
Now, let us find singular value decomposition of A. First we need to change the ring to RDF and find singular value decomposition of this. Now, let us see, suppose we have a unit circle.

(Refer Slide Time: 13:59)



We want to see what happens to a unit circle, when you apply A to it. What happens to the image of unit circle under A?

(Refer Slide Time: 14:08)

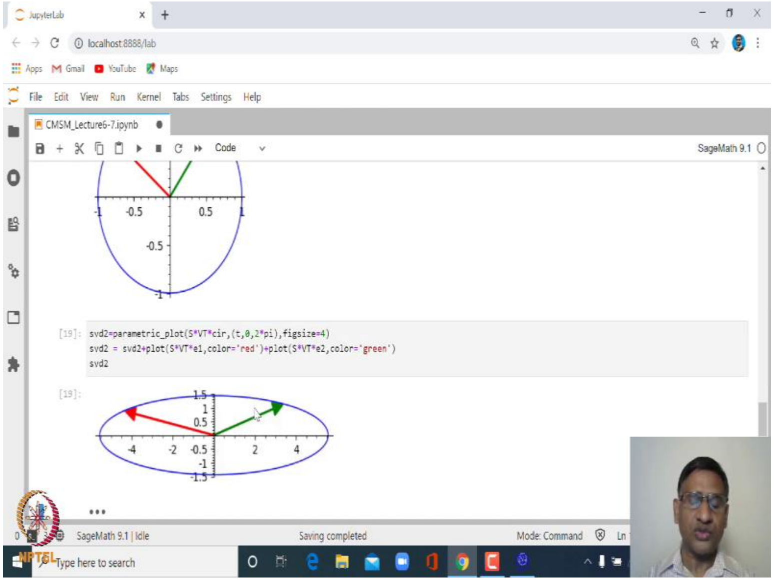


We want to see what happens to a unit circle, when you apply A to it. What happens to the image of unit circle under A.

Let us also plot these unit vectors e1 and e2 and then, let us plot this circle with the image of unit vectors. This is u1, this is e2. When you have multiplied this unit vectors by V transpose, V transpose u1 and V transpose e2, then these vectors.

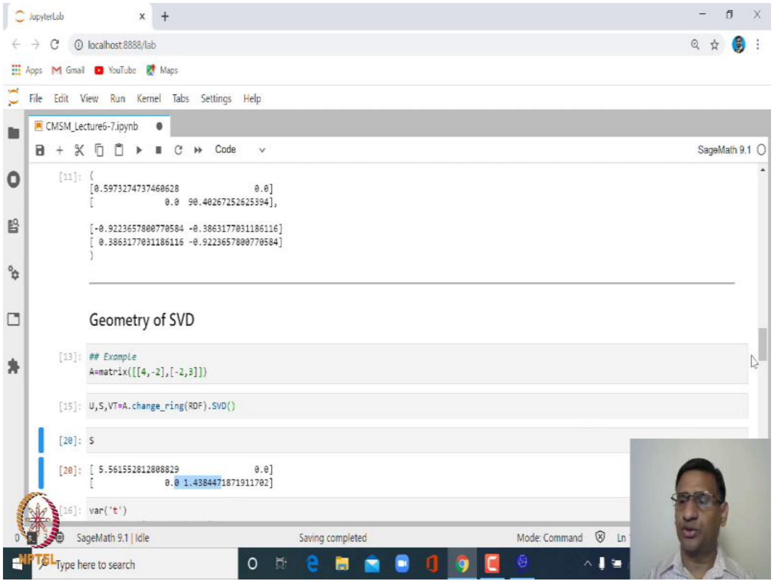
So, for example, this V transpose u1, is in red colour. So, u1 goes to this, and e2 goes to this. Actually you can check that this has been obtained using some kind of rotation.

(Refer Slide Time: 14:58)



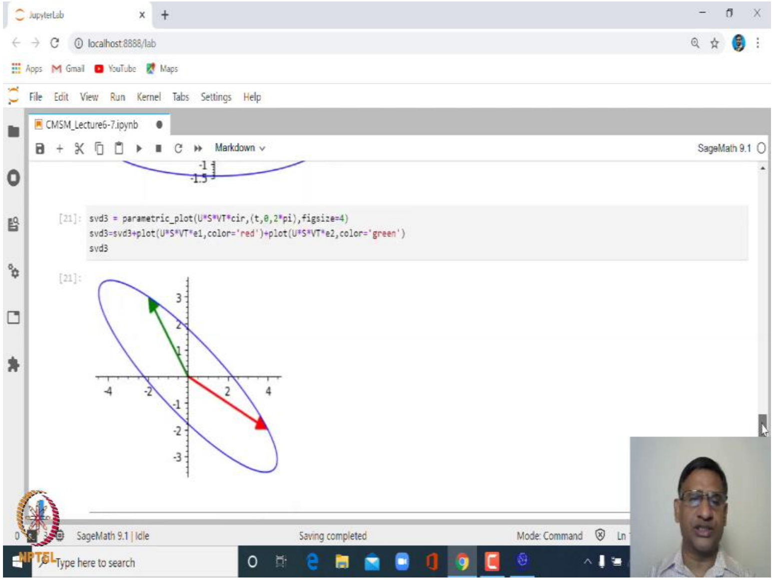
Now, next let us apply sigma to it. We applied V transpose into circle and then S into V transpose the circle. That is how the circle will change and also let us plot the vector S into V transpose e1 and S into V transpose e2. When you plot this, you see what happens. This two e1 and e2 are scaled. So, for example, it is scaled by which factor?

(Refer Slide Time: 15:34)



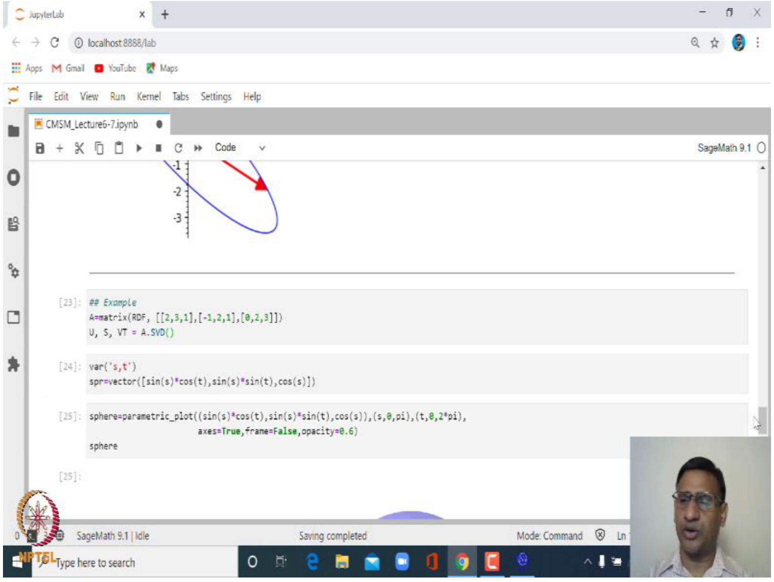
That we can obtained using this S. In this case S is this. So, first x-coordinate will scaled by a factor of 5.56 other co-ordinate is scaled by a factor of 1.43. That is exactly what has happened to these unit vectors right. So, the circle has become a kind of ellipse.

(Refer Slide Time: 16:03)



And now again applied U to it. So, when we apply U to it you will get some kind of rotation. So, U into S into V transpose of circle. Here again to u1 apply A, which is same as saying U into S into V transpose. Similarly, U into S into V transpose of e2. So, that is how you get. So, now this ellipse is rotated like this by U. This is what it means actually. So, when you take any circle and apply A to it, this would become a kind of ellipse, this is the semi-major and semi minor axes, this will get rotated by once by V transpose and other one by U.

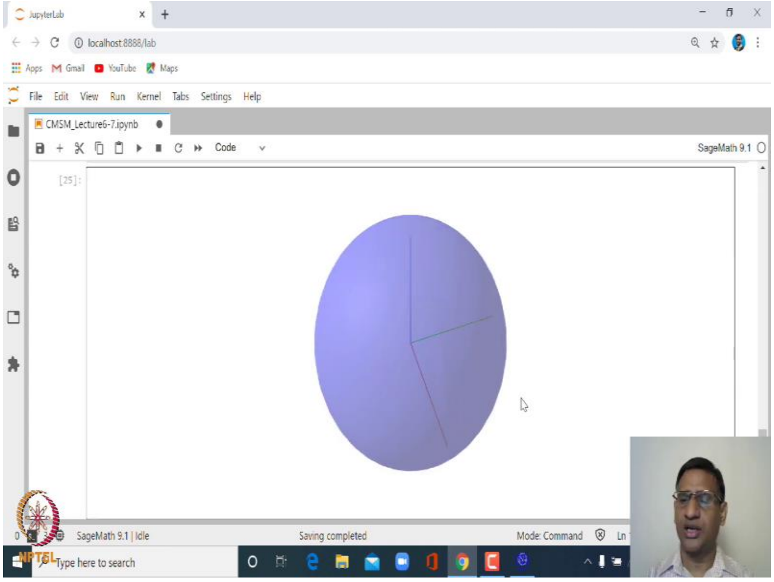
(Refer Slide Time: 16:56)



Similarly, you can do the same thing using 3 cross 3 matrix. Now, let us take a 3 cross 3 matrix 2, 3, 1 minus 1, 2, 1 and 0, 2, 3 and find out again singular value decomposition of A. U S V transpose V T which is V transpose as A into A dot SVD and let us take a sphere.

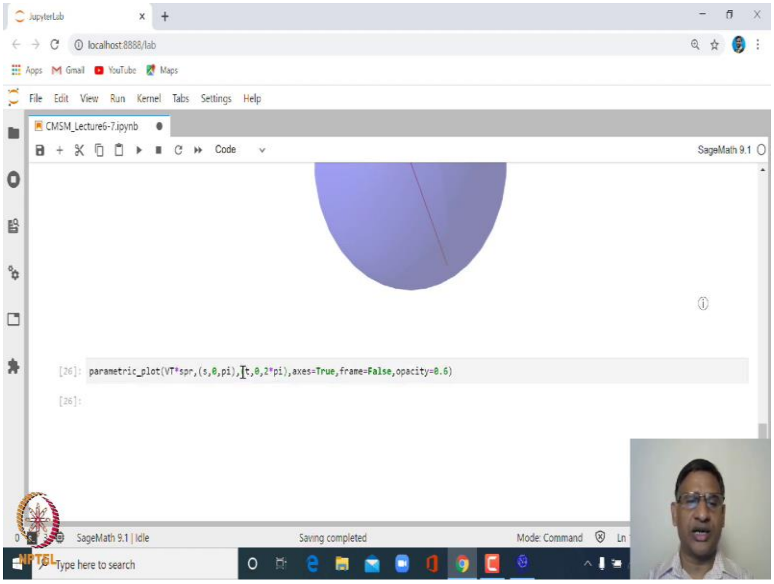
The parametric co-ordinate of a sphere is nothing but the first co-ordinate is sin s cos t second co-ordinate is sin s sin t and third coordinate is cos s. Here, s and t s will vary between 0 to 2 pi t will vary between 0 to 2 pi or other way round ok. So, let us plot. So, actually in this case s varies between 0 to pi and t varies between 0 to 2 pi. So, let us plot this sphere as a parametric plot.

(Refer Slide Time: 17:55)

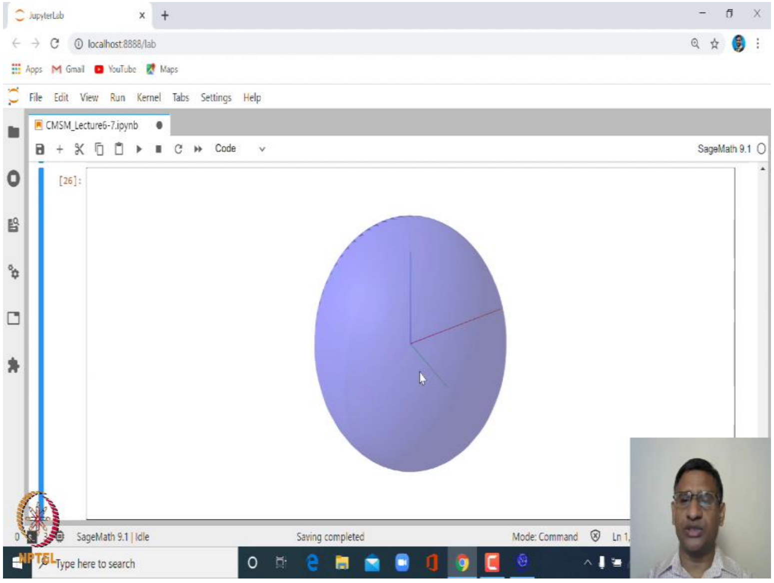


So, this is what you get. This is sphere and you can see the co-ordinate axis here. Now, we want to check what happens to this sphere when we apply A to it. When you multiply this sphere by A what happens. Multiplying A to this sphere is same as saying multiply first by V transpose then, multiply by S and then multiply by U.

(Refer Slide Time: 18:25)



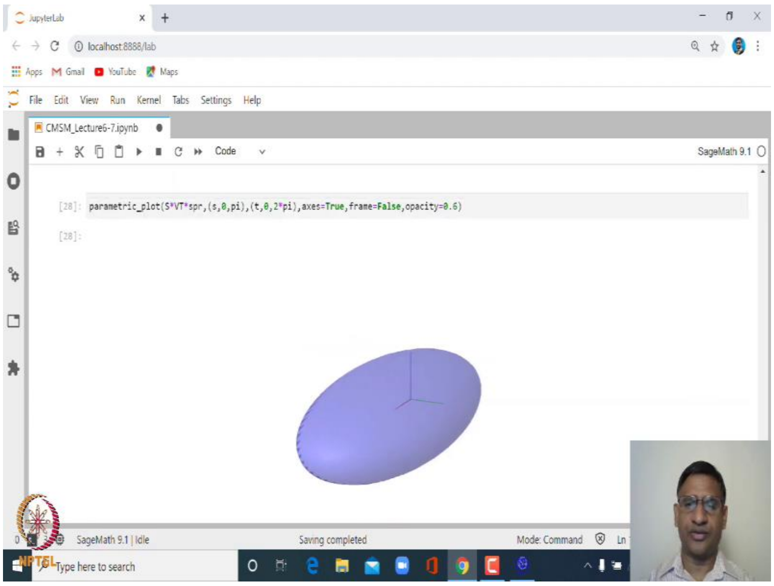
(Refer Slide Time: 18:29)



So, let us sequentially plot a graph of each of this. So, first let us plot graph of sphere multiplied by V transpose. V transpose times sphere. V transpose is actually orthogonal matrix. So, V transpose sphere will just be will get rotated. When you rotate a sphere, what you are going to get is a sphere itself. Of course, we can check what happens to unit vectors as we did in case of circle. That I will leave it as an exercise.

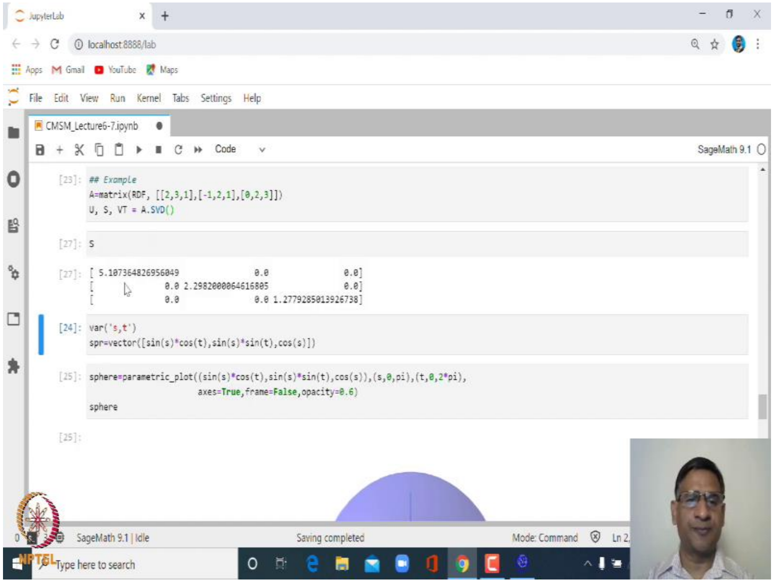
You can try to explore what happened to the unit vectors. Try to plot unit vectors e1, e2, e3 along with the sphere and then add what happens to e1, e2, e3 under V transpose.

(Refer Slide Time: 19:13)



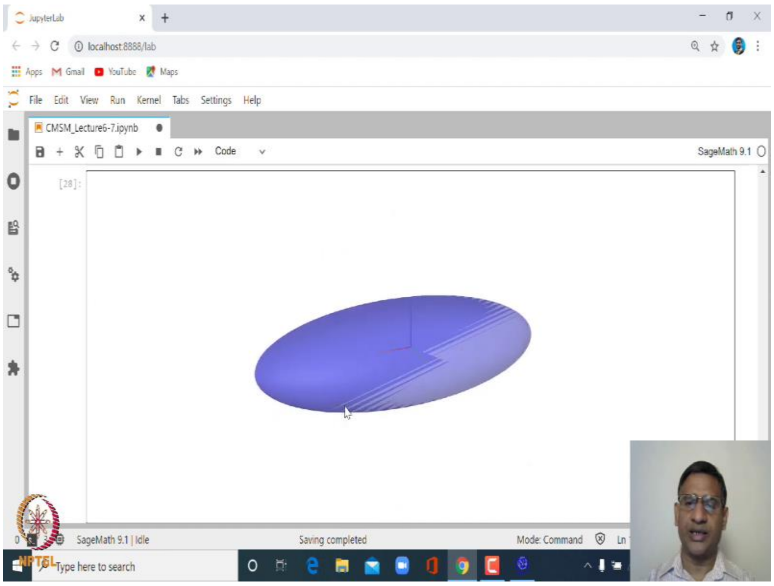
Similarly, next let us apply S to it. V transpose sphere times S. Now, what you expect it to do? Let us just look at what is S? S will be diagonal matrix.

(Refer Slide Time: 19:34)



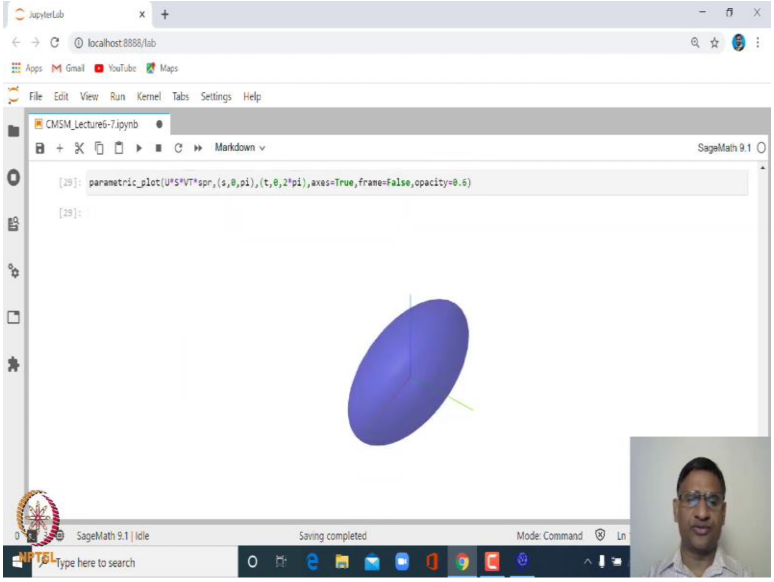
The first axis will get multiplied by 5.107. Second one will by about 2.292 and third one about by 1.0277. So, you this sphere now, will become an ellipsoid. Let us see that. Yes, that is what has happened. So, the sphere has become ellipsoid.

(Refer Slide Time: 20:08)



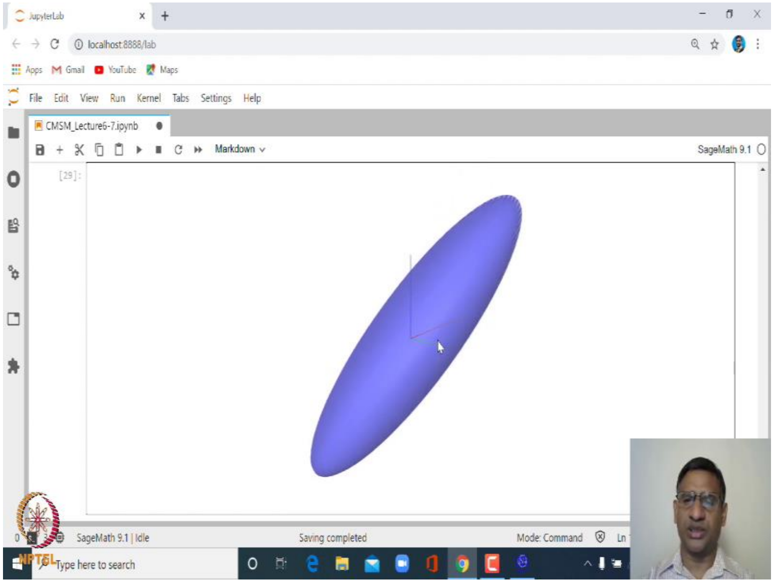
And you can see what happens to this unit vectors along the co-ordinate axis.

(Refer Slide Time: 20:15)



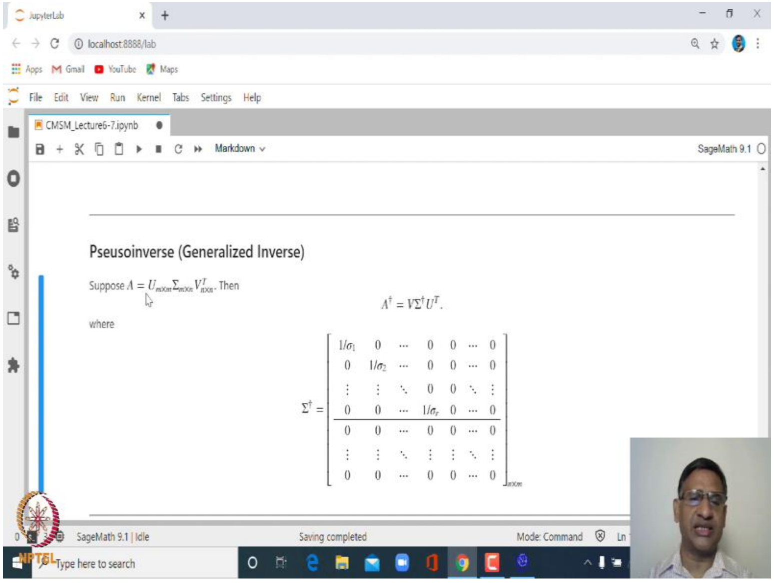
Then apply U to it. So, U times S times V transpose. So, this ellipsoid will get actually rotated by this U, because U is a orthogonal matrix. That is what you can see here.

(Refer Slide Time: 20:33)



Again it just gets rotated right. There is no change in the coordinates axis, but it only gets rotated. That is how you can explain geometric meaning of singular value decomposition, the multiplication of any matrix to any object in R2 and R3 right.

(Refer Slide Time: 21:08)



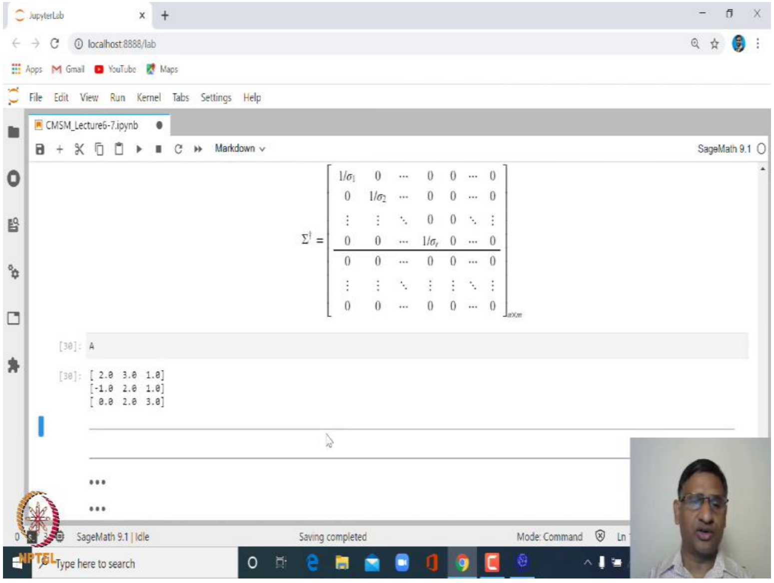
Next let us look at how we can obtain generalized inverse using singular value decomposition. We have we have split A as U into sigma into V transpose. A is m cross n matrix and A is U into sigma into V transpose. You look at what is the property of the inverse of multiplication of two matrices. Inverse of A into B is nothing but B inverse into A inverse.

So, if I take inverse of this multiplication, this is going to inverse of A. Suppose, if it exists inverse of A is going to be inverse of U into sigma into V transpose, which is nothing but inverse of V transpose that is same as V into inverse of sigma and that and then, multiplied by inverse of U, which is U transpose.

But A may not be invertible, U and V are orthogonal matrices, they are invertible. Sigma is not a square matrix, this is a rectangular matrix. So, you have to define what is inverse of this kind of rectangular matrix, what is generalized inverse of this kind of rectangular matrix.

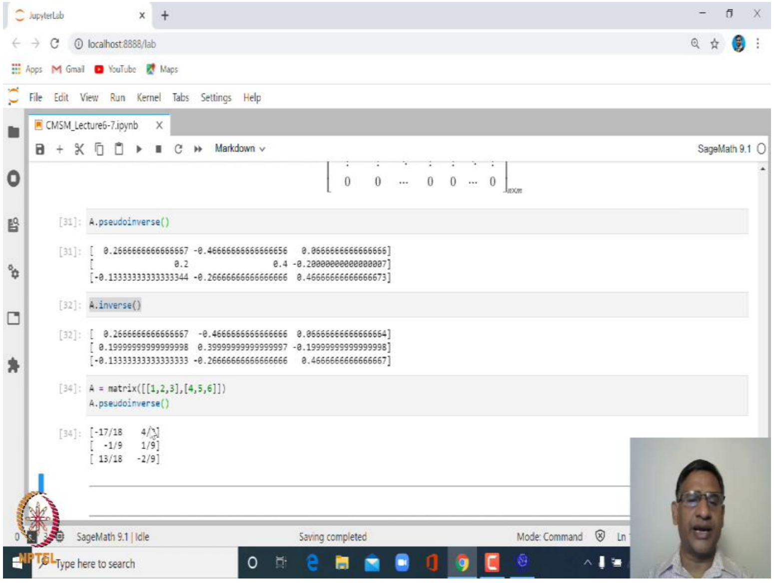
Now, let us just imagine, in case, Sigma happens to be square a diagonal matrix and if diagonal entries are all invertible, its inverse is nothing but the diagonal entries become one upon the diagonal entries of each of these. In case, diagonal entries are zero then, you cannot divide 1 by the diagonal entries. Then the generalised inverse of this sigma is going to be nothing but you just take a transpose of this matrix sigma and wherever you have diagonal entries which is non-zero, you take reciprocal. That is going to be generalized inverse of sigma. So, this is quite easy to check. The generalized inverse of A, we denote it write it as A dagger is going to be V into generalized inverse of sigma into U transpose. So, that is very easy do this computation. Once we know singular value decomposition, this is very easy to compute. In this case all only the concept which is needed was the finding generalized inverse sigma which is a kind of diagonal matrix. So, this is how you can obtain. But of course, sage has inbuilt function to find the generalized inverse.

(Refer Slide Time: 23:50)



For example, take any matrix A. In this case A is a 3 cross 3 matrix.

(Refer Slide Time: 23:59)



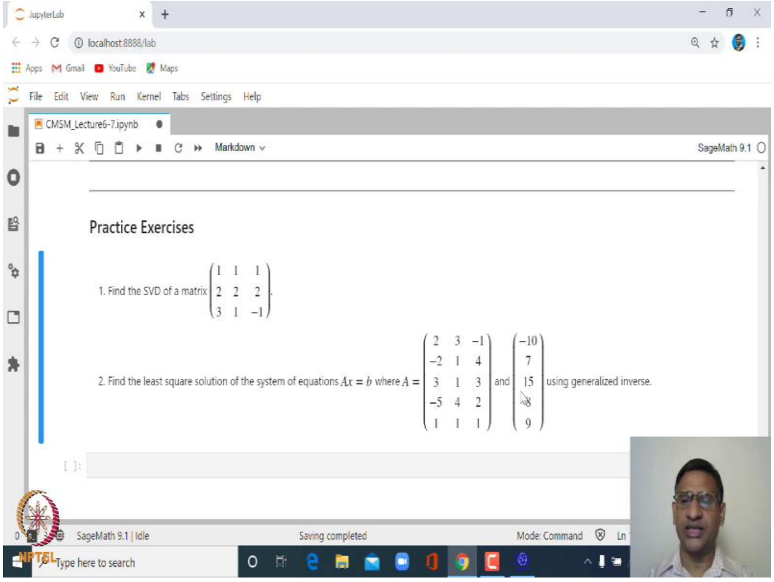
Let us let us define A dot, it is called I think pseudoinverse. So, this is a generalized inverse. Now, if you look at what is A inverse? I this case A dot inverse if it exists. So, both are same.

In case A is invertible pseudo-inverse or generalized inverse will be same as its inverse. You can see here, they differ in decimal. The algorithm to compute pseudo-inverse may be different from the algorithm to compute inverse.

For example, let us take A to be a matrix, let us say, 1 2 3 and second row is let us say, 4 5 6 and we can find its generalized inverse. A dot pseudo-inverse, this is what you get pseudo-inverse of this.

You can try to verify the properties of this pseudo-inverse. For example, compute what happens to A into pseudo-inverse of A, The pseudo-inverse of A into A. All these properties you can you can verify. As I said earlier, using pseudo-inverse, you can also obtain least square solutions.

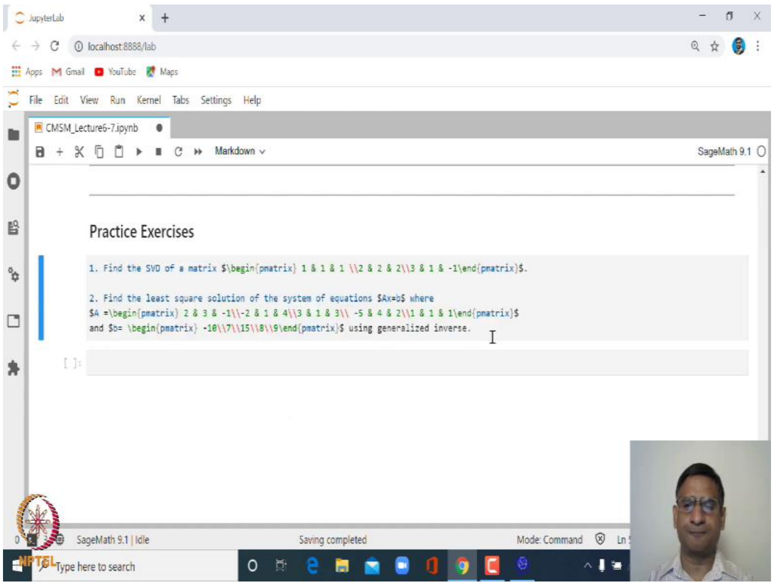
(Refer Slide Time: 25:31)



Let me leave few exercises for you. First is to find singular value decomposition of this matrix. Quite simple!

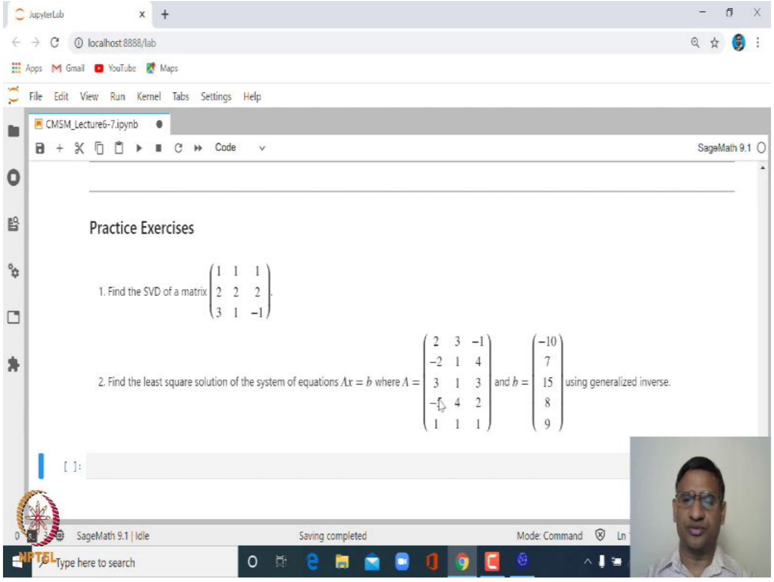
The second one is find least squares solution of Ax equal to b, where A is this.

(Refer Slide Time: 25:46)



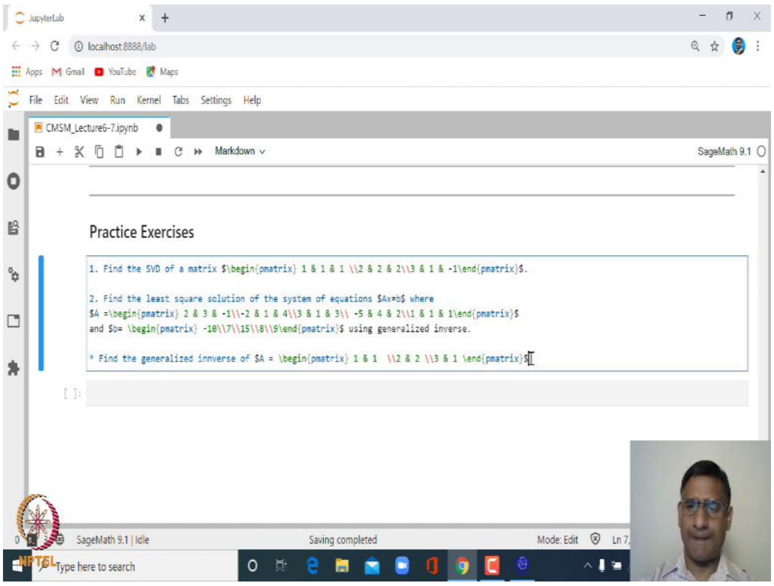
And b is this, using pseudo-inverse or using generalized inverse.

(Refer Slide Time: 25:50)



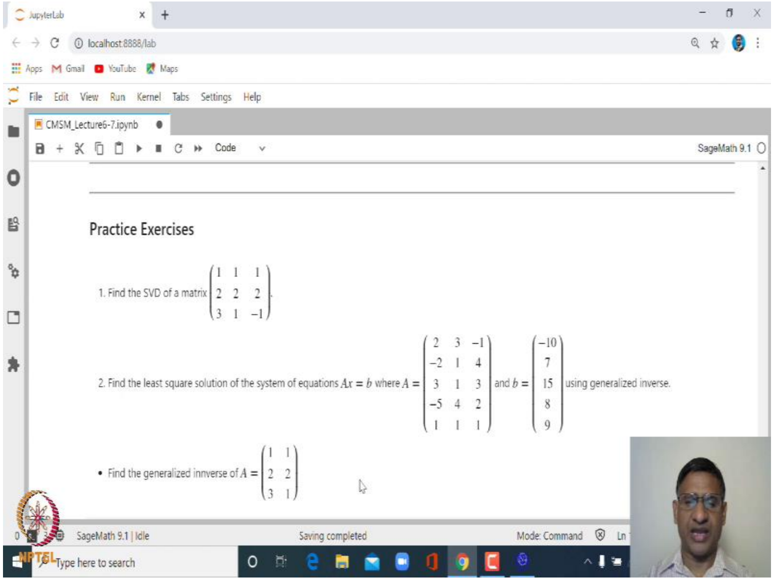
The least squares solution, x star is going to be a pseudo-inverse of A times b. So, you simply compute pseudo-inverse of this, which is we denote by A dagger. So, A dagger times b is what you need to compute. It is a very easy exercise. Now, let me leave one more exercise.

(Refer Slide Time: 26:19)

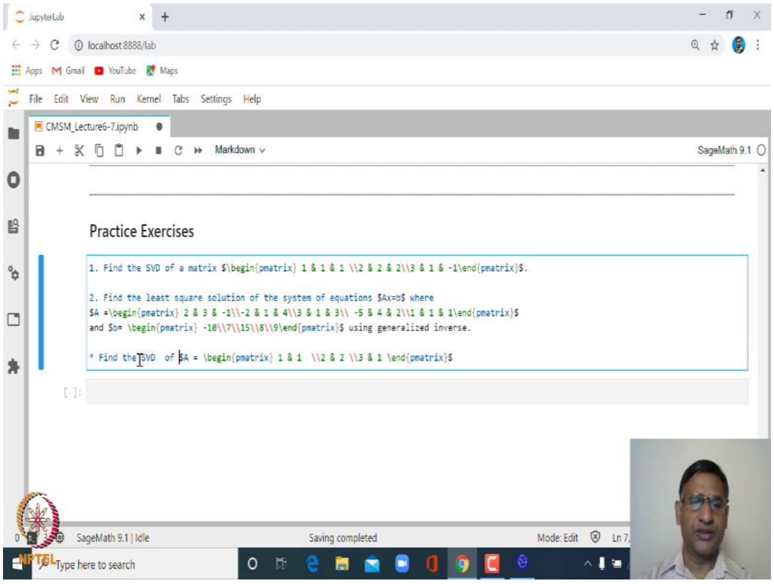


Find the generalized inverse of let us say A, which is, let me just copy this here, A is equal to this 3 cross 2 matrix. Earlier rows what was the one which we took was 2 cross 3 matrices. So, in this case now, what you may have to do find the generalized inverse.

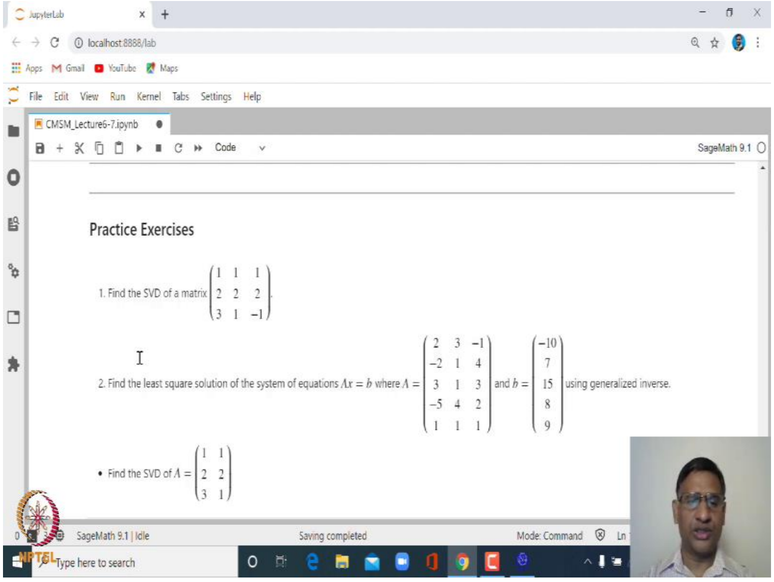
(Refer Slide Time: 27:05)



(Refer Slide Time: 27:28)



(Refer Slide Time: 27:32)



I will say SVD of this matrix. So, in this case what you may have to do. When you look at this A transpose A, A A transpose in this case is going to be a 3 by 3 matrix whereas, A transpose A will be 2 cross 2 matrix. So, you need to define two vectors v1 and v2 for a capital V and then, three vectors u1, u2, u3 for U.

Now, the question is how we obtain; u1 and u2, you can obtain from v1 and v2, but for u3 you may have to take a vector which is orthogonal to u1 and u2; that is what it means. That is why I gave you this exercise.

We will look at some more applications in the next class. So, let me stop here.

Thank you very much.