Matrix Analysis with Applications
Dr. Sanjeev Kumar
Department of Mathematics
Indian Institute of Technology, Roorkee

Lecture - 29
Pseudo-Inverse and SVD

Hello friends. So, welcome to the lecture on Pseudo-Inverse and Singular Value Decomposition. In the last lecture we have learned least square approximation of over determined and under determined systems. However, we have used the matrix A transpose into a in case of over determined system and A into A transpose in case of under determined system.

And, in particular for finding the pseudo inverse we have use the inverse of these 2 matrices. At that point I told you that if the inverse does not exist, then we will discuss this case later. So, in this lecture we will discuss that case.

(Refer Slide Time: 01:14)

Let $AX = b$ be an overdetermined system, where A is a $m \times n$ matrix with $m > n$.

$$X = A^+ b \quad , \text{where } A^+ = (A^TA)^{-1} A^T$$

If $\text{rank}(A) < n \Rightarrow |A^TA| = 0$ and hence $(A^TA)^{-1}$ does not exist.

Let $A = USV^T$ be the SVD of matrix A, where U is a $m \times m$ orthogonal matrix and V is a $n \times n$ orthogonal matrix and the matrix S is a $m \times n$ matrix containing the singular values of A.

$$AX = b \Rightarrow (USV^T)X = b$$

$$X = (USV^T)^{-1} b = ((V^T)^{-1} S^+ U^{-1}) b$$

$$\Rightarrow \quad X = (V\underline{S^+}U^T) b$$

So, let A X equals to b be an over determined system, where A is a m by n matrix with m greater than equals to n because it is an over determined system. So, in the last lecture we have learned the least square approximation of this can be written as like this X equals to A plus into b, where A plus is the pseudo inverse of A and it is defined as A transpose A inverse into A transpose. So, if rank of the matrix A is less than n. This implies that A transpose A and the determinate of A transpose A equals to 0 and hence A transpose A inverse does not exist.

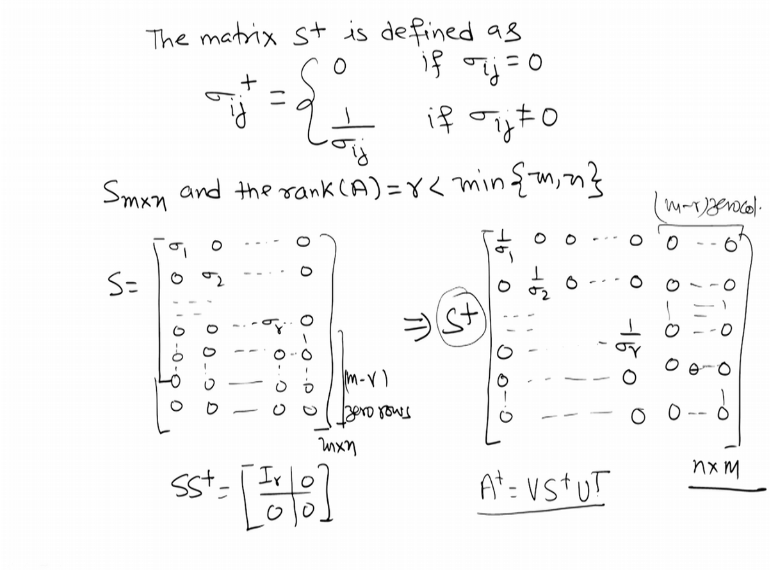
So, now question is how to find least square approximation of such a system?. So, for doing this we will use make use of singular value decomposition for calculating the pseudo inverse in these this particular case. So, let A equals to USV transpose be the singular value decomposition of matrix A, where U is a m by m orthogonal matrix and V is a n by n orthogonal matrix and the matrix S is A m by n matrix containing the singular values of A. So, means we are using here full singular value decomposition.

So, now, we are having a system A X equals to b. So, a can be written as USV transpose into X equals to b from here I can write X equals to USV transpose inverse into b. This can be written as V transpose inverse let us say S inverse I am writing as S plus into U inverse into b. So, from here I can write X equals to since V is an orthogonal matrix. So, V transpose is also orthogonal. And here V transpose inverse will become V transpose transpose which will be your matrix V.

S plus and U inverse will become U transpose because U is also an orthogonal matrix into b. Now, you can easily write V and U transpose, because this you are having already in singular value decomposition of A.

Now, the question arise how to calculate this S plus?

(Refer Slide Time: 06:06)



For the matrix S plus is defined as. So, if sigma i j are the entries of S, then the entries of S plus I am defining as sigma ij plus and this equals to 0 if sigma ij equals to 0 and it is equals to 1 upon sigma ij if sigma ij is non zero.

So, for example, if you are having A matrix S, which is m by n and the rank of the matrix A is r r is less than minimum of m and n. So, it will be less than m in case of over determined under determined system and less than n in case of over determined system. So, in both the case A transpose A as well as A into A transpose will be singular matrices.

And, hence the inverse of these 2 matrices do not exist. So, in this case my S will be something of this shape sigma 1 0 0 0 sigma 2 0 and then I will be having 0 0 sigma r 0 0 00. So, these will be matrix A in case of over determined system A m by m by n matrix and then what I am having. So, it is m by n matrix. So, these will be total m minus r 0 rows.

In this case your S plus will become A n by m matrix now which is having 1 upon sigma 1 because sigma 1 is non-zero here. So, it will become 1 upon sigma 1 0 0 0 0 1 upon sigma 2 0, in the same way it will become one upon sigma r and then what I will be having; I will be having n minus m 0 rows, because there will be 1 upon 0 so, 1 upon 0 I will replace by 0 we will see this in an example and then so, this will become n by m matrix which is having now m minus r 0 columns, columns if you see here S into S plus will become an identity matrix of r by r and then 0 blocks accordingly.

So, in the first r rows there will be one in the as the diagonal entries and rest of the places it will be 0. So, hence in this way we can define the pseudo inverse of S as S plus. And, the pseudo inverse of A plus A will become A plus which is V S plus into U T. So, let us take an example of this.

(Refer Slide Time: 10:44)

Ex:- Consider an example of fitting a line through data points

$(1, 2)$, $(2, 3)$ and $(3, 5)$
$x_1\ y_1 \quad x_2\ y_2 \quad x_3\ y_3$

In Matlab
$[U\ S\ V] = SVD(A)$

Sol^n: Let the equation of line becomes

$$y = mx + c$$

$$\begin{bmatrix} x_1 & 1 \\ x_2 & 1 \\ x_3 & 1 \end{bmatrix} \begin{bmatrix} m \\ c \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} \Rightarrow A = \underbrace{\begin{bmatrix} 1 & 1 \\ 2 & 1 \\ 3 & 1 \end{bmatrix}}_{\substack{3\times 2 \\ m\times n}} ; b = \begin{bmatrix} 2 \\ 3 \\ 5 \end{bmatrix}$$

$$A \quad X = b$$

SVD of $A = USV^T$, where

$$U = \begin{bmatrix} -.3231 & .8538 & .4082 \\ -.5475 & .1832 & -.8165 \\ -.7719 & -.4873 & .4082 \end{bmatrix} ; S = \begin{bmatrix} 4.0791 & 0 \\ 0 & .6005 \\ 0 & 0 \end{bmatrix} ; V = \begin{bmatrix} -.9153 & -.4027 \\ -.4027 & .9153 \end{bmatrix}$$

$$A^+ = VS^+U^T = V \begin{bmatrix} \frac{1}{4.0791} & 0 & 0 \\ 0 & \frac{1}{.6005} & 0 \end{bmatrix} U^T = \begin{bmatrix} -.5 & 0 & .5 \\ 1.3333 & .3333 & -.6666 \end{bmatrix}$$

Consider an example of fitting a line through data points let us say 1 2, 2 3, and 3 5. So, let us solve this example using the approach which I told you just now. So, here let the equation of line becomes y equals to mx plus c. So, here I am having x 1 y 1 x 2 y 2 x 3 y 3.

So, now from the first equation I will get or in matrix form I can write it x 1 1 x 2 1 x 3 1 with multiplied with mc equals to y 1 y 2 y 3. So, please note that in each case the coefficient of c will be one, because here c is having coefficient as one. So, if this is my matrix A this is X this equals to b, then my matrix A will become here 1 1 2 1 and 3 1, x 1 is 1 x 2 is 2 x 3 is 3 and here b will become that is the right hand side vector y 1 y 2 and y 3.

So, now, I need to find out the values of m and c. So, it is an over determined system if I perform the S V D of A then the matrix A can be written as USVD USV transpose. So, where the matrix U will be A 3 by 3 matrix because m is 3 here it is 3 by 2. So, m by n and the entries of this will be minus 0.3231 0.8538 0.4082 minus 0.5475 0.1832 and minus 0.8165 will be the second row and third row will be minus 0.7 7 1 9 minus 0.4 8 7 3 and 0.4 0 8 2. Here, the matrix S will become A 3 by 2 matrix, which will be having 4.0791 is the first singular value of A the biggest 1 0 then 0 sigma 2 will be 0.6005 and then A 0 row m minus n number of 0 row. So, this will be 3 minus 2 1 1 0 row. And finally, V will be A 2 by 2 matrix the entries of V will be minus 0.9153 minus 0.4027 minus 0.4027 and 0.9153 here U and V are orthogonal matrices.

And, we have learn in the previous lecture that how to calculate these matrices? The alternate way of doing it is using the MATLAB software and there you are having direct command for finding the singular value decomposition edge SVD of A. So, if you will perform U S V equals to SVD of A in MATLAB, you will get these 3 matrices. So, this is just an alternate for doing it to save the calculation efforts.

So, I am having now these 3 matrices. So, now, my pseudo inverse of a will become A plus, which will be V S plus U T here V will be this matrix S plus will be now A 2 by 3 matrix. So, it will be 1 upon 4.0791 000 1 upon 0.6 0 0 5 0 into U transpose. And, if I calculate these matrix A plus so, it will be A now 2 by 3 matrix, because a is 3 by 2 matrix which is given as minus 0.50 0.5 1.3333 0.3333 and minus 0.6666. So, this is my A plus.

(Refer Slide Time: 17:30)

$$X = A^+ b = A^+ \begin{bmatrix} 2 \\ 3 \\ 5 \end{bmatrix} = \begin{bmatrix} 1.5 \\ 0.3333 \end{bmatrix} \quad \text{Ans.}$$

LSA with SVD (LLS)

Consider $Ax = b$, where A is $m \times n$ matrix with $\text{rank}(A) = r < \min\{m, n\}$

then

$$\min_{x} \|Ax - b\|_2^2$$

So, now if I calculate my X, which will become A plus into b. So, A plus into my matrix b is 2 3 5 the vector b. So, which comes out to be 1.5 and 0.3333 as the least square approximation of this system so, in this way we can calculate the least square approximation solution using the singular value decomposition. In particular we have used the singular value decomposition for calculating the pseudo inverse.

Now, let us make the analysis of it how we have done it L S A with SVD or this L S A I can also write L L S means linear least square solution. So, for doing this again consider AX equals to b, where A is m by n matrix with rank of A equals to r, which is less than minimum of m n. Then, if it is a an over determined system, the least square solution is 2 minimize square of A X minus b A X norm of this and we have to find out such X.

(Refer Slide Time: 19:44)

$$X = A^{+} b = A^{+}\begin{bmatrix} 2 \\ 3 \\ 5 \end{bmatrix} = \begin{bmatrix} 1.5 \\ 0.3333 \end{bmatrix} \quad \underline{\text{Ans.}}$$

LSA with SVD (LLS) $\qquad \|UX\| = \|X\|$

Consider $AX = b$, where A is $m \times n$ matrix with $\text{rank}(A) = r < \min\{m, n\}$; $A = \underline{U}S\underline{V}^T \Rightarrow S = U^T A V$

then

$$\|AX - b\|_2^2 = \|U^T(A\underbrace{VV^T}_{I}X - b)\|_2^2$$

$$= \|(U^T A V)(V^T X) - U^T b\|_2^2$$

$$= \|S V^T X - U^T b\|_2^2 \qquad \text{Let } V^T X = Z$$

$$= \|SZ - U^T b\|_2^2$$

$$= \sum_{i=1}^{r} (\sigma_i z_i - u_i^T b)^2 + \sum_{i=r+1}^{m} (u_i^T b)^2$$

So, if in particular I check about this particular thing then, this I can write as U transpose A V V transpose X minus b, here you know that V into V transpose is I. So, A V V transpose X will become simply A into X minus b. So, this bracket term is similar to this one. Moreover I am pre multiplying this particular vector by U transpose. And, in singular value decomposition of A you know that U end V are the orthogonal matrices.

Since, U is orthogonal transformations. So, U transpose is also an orthogonal transformation and we know that if U is orthogonal, then U X equals to norm of U X equals to norm of X; means norm preserve under the orthogonal matrix multiplication. So, the same thing I am doing here now this thing I can write U T A V into V T X minus U transpose into b. Now, A equals to U S V T. So, from here I can write S equals to U transpose A into V.

So, this I am replacing with S. So, S V T X minus U T b square; so this equals to this one let V transpose X equals to Z another vector. So, from here what I can write this can be written as S into Z minus U T b this one or if I open this norm this can written as I equals to 1 to r sigma I z I minus u i T b, where U I are the rows of matrix U square and why I am taking up to r because the matrix A is having only r nonzero singular values. After, that all sigma I where I is greater than r will become 0 so, this term will become 0. So, if this will become 0 then I equals to r plus 1 to m it is will become simply this.

(Refer Slide Time: 23:02)

$$z_i = \begin{cases} \dfrac{u_i^T b}{\sigma_i} & ; \quad i = 1, 2, \cdots, r \\ \text{arbitrary} & \text{for } i = r+1, \cdots, n \end{cases}$$

As a result

$$\min_{x} \|Ax - b\|_2^2 = \sum_{i=r+1}^{m} (u_i^T b)^2$$

Recall $z = V^T x \Rightarrow x = Vz$

$$\|x\|_2 = \|V\underline{V^T x}\|_2 = \|Vz\|_2 = \|z\|_2$$

Hence, LSA of $Ax = b$ given as

$$\boxed{x^* = \sum_{i=1}^{r} \frac{u_i^T b}{\sigma_i} v_i}$$

So, from here I can write z i equals to u i transpose b upon sigma i for i equals to 1 2 up to r, from the first term and then it will be some arbitrary for i equals to r plus 1 up to n. As, a result if z i equals to this for first r, then first term will become 0. So, minimum X and the least square error in this solution will become summation i equals to r plus 1 to m u i T b whole square. Also, you know that recall Z equals to V T X. So, from here I can write X equals to V into Z, because V is an orthogonal matrix.

Moreover, the length of the vector X, equals to V into V transpose into X, because V into V transpose is an identity matrix this will become V T X is z. So, V into z and since V is an orthogonal operator. So, it will preserve the norm this equals to z. So, from here I can write hence least square approximation of A X equals to b given as X star that is the solution I equals to 1 to r u I T b upon sigma i this is my z i and then into v i, because z equals to V transpose X.

So, in terms of singular values I can write the solution of least square solution in this way. So, if we see the earlier example, which we have taken for fitting a line.

(Refer Slide Time: 25:53)

$$X^* = \sum_{i=1}^{2} \frac{u_i^T b}{\sigma_i} \vartheta_i = \frac{u_1^T b}{\sigma_1} \vartheta_1 + \frac{u_2^T b}{\sigma_2} \vartheta_2$$

$$= -1.5072 \begin{pmatrix} -.9153 \\ -.4027 \end{pmatrix} - .298 \begin{pmatrix} -.4027 \\ .9153 \end{pmatrix}$$

$$= \begin{pmatrix} 1.4995 \\ 0.33342 \end{pmatrix} \approx \begin{pmatrix} 1.5 \\ 0.333 \end{pmatrix} \qquad \cancel{A^{\pm}} \quad \underline{X = A^+ b}$$

There X star that is the least square solution can be given as i equals to 1 to 2 because we are having only 2 singular values u i T b upon sigma I v i. So, this will become u 1 transpose into b upon sigma 1 into v 1 plus u 2 transpose b upon sigma 2 into v 2. And, if I calculate this particular term it is a scalar value this comes out to be 1.5072 and then v 1 is now the first column minus 0.9153 and this will become 4027 minus 0.298 this is this particular scalar value, that is minus of 0.298 multiplied with second column of the v that is minus 0.4027 and then 0.9153.

So, if I do it comes out to be 1.4995 and then 0.33342, which is the same 1.5 and 0.333, which we obtained earlier with A plus equals to means which we obtained already with X equals to A plus into b hence the claim is verified. So, this is another way of doing the analysis of least square approximation using the singular value decomposition. And, here we have seen that the both the answers are equal.

Now, take one more thing means one more example where we are having 0 determinant. So, take one more example where we are having 0 determinant of the matrix A.

(Refer Slide Time: 28:19)

Ex.2 Solve

$$\left.\begin{array}{l} x_1 - 2x_2 + x_3 = 3 \\ 2x_1 - 4x_2 \qquad = 0 \\ x_1 - 2x_2 + 3x_3 = 9 \end{array}\right\} \quad \& \quad A = \begin{bmatrix} 1 & -2 & 1 \\ 2 & -4 & 0 \\ 1 & -2 & 3 \end{bmatrix};\ b = \begin{bmatrix} 3 \\ 0 \\ 9 \end{bmatrix}$$

$$\det(A) = 0$$

$A = USV^T$ where

$$U = \begin{bmatrix} -.4237 & .0531 & -.9045 \\ -.7227 & -.6219 & .3015 \\ -.5465 & .7813 & .3015 \end{bmatrix};\ S = \begin{bmatrix} 5.7807 & 0 & 0 \\ 0 & 2.5659 & 0 \\ 0 & 0 & 0 \end{bmatrix};$$

$$V = \begin{bmatrix} -.4178 & -.1596 & .8944 \\ -.8355 & .3192 & .4472 \\ -.3568 & .9342 & 0 \end{bmatrix}$$

$$A^+ = VS^+U^T = V \begin{bmatrix} \frac{1}{5.7807} & 0 & 0 \\ 0 & \frac{1}{2.5659} & 0 \\ 0 & 0 & 0 \end{bmatrix} U^T = \begin{bmatrix} \quad & \quad & \quad \end{bmatrix}_{3\times 3}$$

$$X = A^+ b = \begin{bmatrix} 0 \\ 0 \\ 3 \end{bmatrix}$$

So, example 2 solve x 1 minus 2 x 2 plus x 3 equals to 3 2 x 1 minus 4 x 2 equals to 0, and then x 1 minus 2 x 2 plus 3 x 3 equals to 9 and solve means find out the minimum norm solution in least square approximation of this system.

Here, if you see the matrix A is given as 1 minus 2 1 2 minus 4 0 1 minus 2 3 and the right hand side vector b is 3 0 9. If, you check here determinant of A comes out to be 0 and hence you cannot obtain A inverse and you cannot find the exact solution like X equals to A inverse b. So, here if I perform the singular value decomposition of A, then A will become U S V transpose where U is again will be A 3 by 3 matrix. So, it will I am writing it.

So, this is my matrix U the matrix S will be A diagonal matrix A 3 by 3 diagonal matrix having singular values of A. So, first singular value is 0 0 0 second singular value is 2.5 6 5 9 0 and then 0 0 third singular value is 0, because determinant is 0. So, at least 1 of the singular value will be 0.

And finally, matrix V is given as 4 1 7 8 minus 0.1596 0.8944 0.835 5.3192 0.4472 and then minus 0.3568 0.9342 and this entry 0. So, here pseudo inverse of A plus will become V S plus U transpose. So, you can write V from here, S plus will become one upon 5.7807 000 1 upon 2.5659000 and according to rule it should be 1 upon 0, but 1 upon 0 is not defined. So, as I told you for writing the S plus 1 upon 0 will be replaced by 0 into U T and this matrix will be A 3 by 3 matrix and then the solution is given as X equals to A plus b, which comes out to be 0 0 3. So, this is the minimum norm solution that is the least square approximation of this system using this approach.

(Refer Slide Time: 32:40)

$$X^* = \sum_{i=1}^{2} \frac{u_i^T b}{\sigma_i} v_i = \left(\frac{u_1^T b}{\sigma_1}\right) v_1 + \left(\frac{u_2^T b}{\sigma_2}\right) v_2$$

$$= \begin{pmatrix} 0 \\ 0 \\ 3 \end{pmatrix} \quad \text{Same as } X = A^+ b = \cdots \quad //$$

If we go with another approach, then the solution X star is given by I equals to 1 to 2, because we are having only 2 nonzero singular values u i transpose b upon sigma i into v i. So, this becomes minus 1 point. So, this will be v 1 sorry not v 1 u 1 transpose b upon sigma 1 into v 1 plus u 2 transpose b upon sigma 2 into v 2.

And, if I calculate this particular term this will be a scalar multiplied with first column of v, this will be again a scalar multiplied with second column of v, this again comes out to be 0 0 3, which is same as X equals to A plus b using the earlier method. So, in this way I have told you the 2 different way of solving linear systems in particular finding the least square solution of the linear systems, in case when the matrix A is having rank r and the size of A is m by n and r is less than minimum of m and n.

So, in the next lecture we will learn another type of systems those are called ill condition systems. And then we will learn how to solve those systems using the concept of singular values.

Thank you very much.