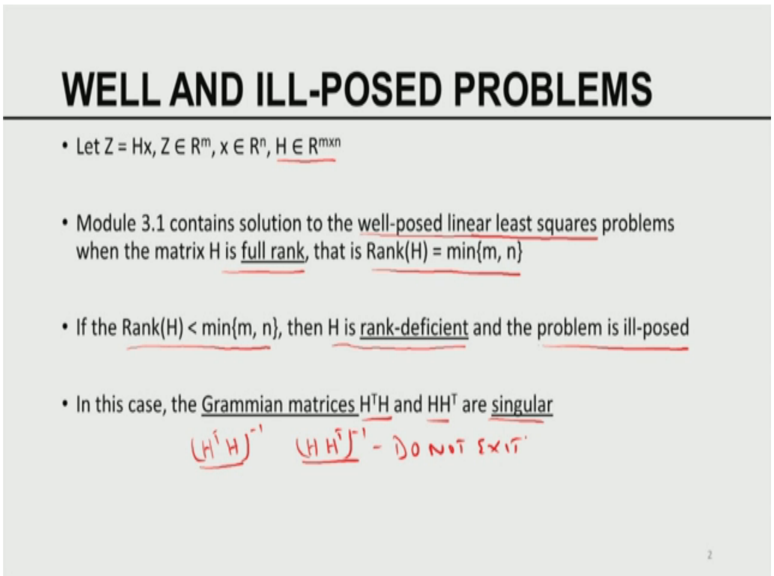
Dynamic Data Assimilation
Prof. S. Lakshmivarahan
School of Computer Science
Indian Institute of Technology, Madras

Lecture - 10
Deterministic, Static, linear Inverse (Ill-posed) Problems

In the last lecture we talked about static deterministic linear inverse problem which are well posed. Why it is well posed? Because we did not find any difficulty in the solution process, that is largely because the matrix H is well is full rank. Because the matrix H is full rank H transpose H is symmetric (Refer Time: 00:39) definite, HH transpose, H transpose H both of them are full rank matrix therefore, we could define the generalized inverse of H both in under determined case and over determined case. We did not have any problem we sailed very smoothly by solving the normal equations.

So, it makes sense to ask our self of the questions what happens if H is such that is not a full rank and that corresponds to ill posed problem. So, ill posed version of the static deterministic linear inverse problem is what we are going to be looking at and we will also talk about what called imperfect model along the way.

(Refer Slide Time: 01:34)

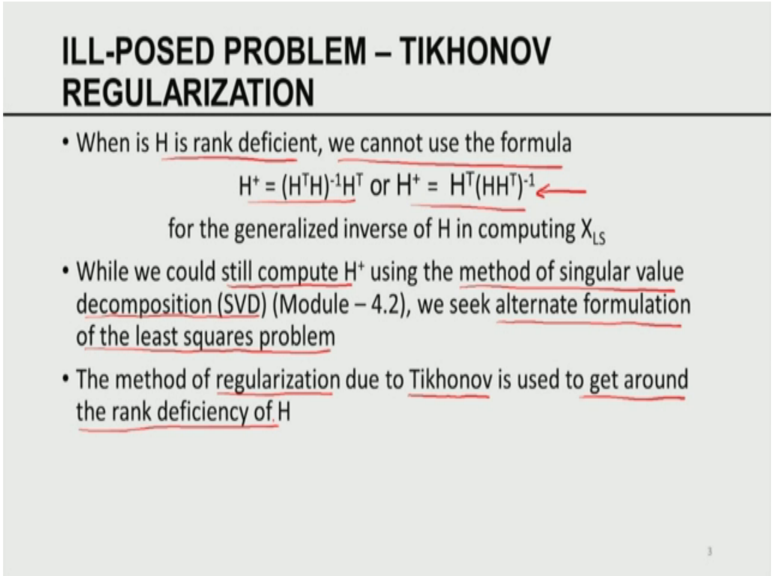


So, Z is equal to H of x is the linear least square problem. H is the matrix which is m by n. In module 3.1 we already talked about the well posed problem linear least squares, the well posed problem essentially banks on the assumption that H is a full rank; that means, the rank of H is minimum of m and n. So, when the over determined case the rank of H is n in the under determined case the rank of H is m. So, we have considered both the cases under one rule namely the rule of H being a matrix of full rank.

Now, we are considering a complimentary case where what happens when H is rank in deficient. So, what does it mean? When the rank of H is not equal to the minimum of m and n, but is less than the minimum of m and n. So, in the over determined case it is not n, but less than n in the under determined case it is not m, but less than m. Such problems are called ill posed problems.

So, well posed verses ill posed largely determined by the properties of the matrix H. Please remember H Z is equal to H of x defines the static model, the properties of model defined by the matrix H, so we considered one aspect of the properties of the model now we are considering another aspect of the properties of the model full rank verses rank deficient. When the H matrix is rank deficient the Grammian matrices H transpose H and H in H transpose are symmetric, but they are not (Refer Time: 03:23) definite, but they are singular. That means, their determines are 0; that means, at least one of their eigen value is 0. If at least one of the eigen value is 0 it is not positive definite it is only positive semi definite. So, positive semi definiteness of H transpose of H and HH transpose leads to singularity and that leads to the fact I cannot simply now compute H transpose, H inverse are HH transpose inverse I cannot compute them because they are all singular, I cannot do, I cannot compute these inverses readily, these do not exist. So, when these do not exit I cannot follow the principles that we have developed in the last lecture. So, this calls for newer methods to be able to handle and we are going to be talking about some of these newer techniques.

(Refer Slide Time: 04:21)

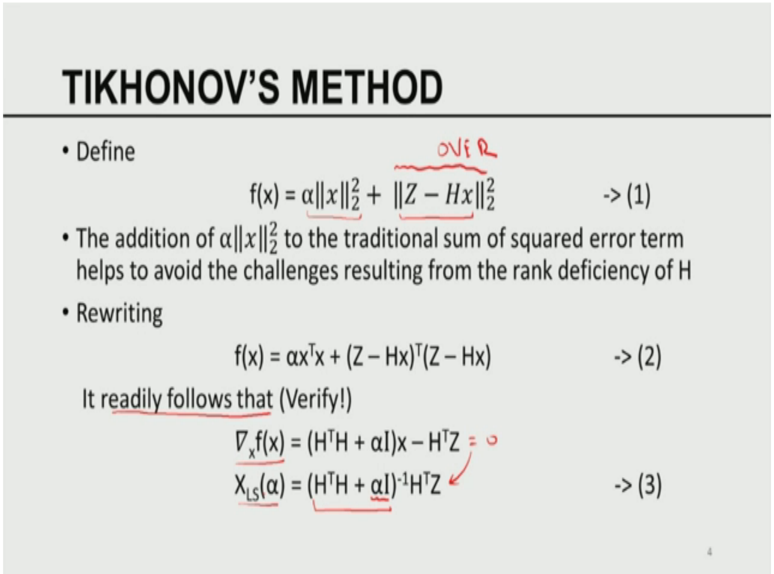


To reinforce further when H is rank deficient when H is rank deficient we cannot use the formula for the generalized inverse of H, this is the generalized inverse of H for the over determined case, this is the generalized formula for the generalized inverse for the over determine case, this is the formula for the under determined case. These generalized inverses that are used in computing X LS now cannot be done, now cannot be computed because they cannot be computed I cannot use the old path way to compute the solution. But the serious generalized inverse still tells you while I cannot compute the generalized inverse using these formulas there is other ways of computing H plus, I can still compute H plus, but using in the method called singular value decomposition.

We will talk lot more about singular value decomposition in one of the modules coming later and using this one can solve, even though one can solve using singular value decomposition at this junctor we are not going to use this we are simply looking for an another alternate formulation of linear least square problems. So, let me clarify where we are. We have a linear least square problems Z is equal to H of x, H is rank deficient I cannot use the method that I have already described thus for, rank deficient cases can be handle by one of the methods called singular value decomposition, we still have not developed the method of singular value decomposition now we will do it later. So, while we will have occasion to revisit the solution of this problem at the time when we describe the singular value decomposition, at this time I am still interested in solving this. So, I am seeking an alternate method, alternate to singular value decomposition. That method is called method of regularization.

This method of regularization was introduced by a Russian mathematician called Tikhonov. This method is meant to get around the rank deficiency of H, is a very simple elegant method to approximate solutions of ill posed problems, rank deficient problems, ill posed rank deficient I am using it synonymously.

(Refer Slide Time: 07:13)



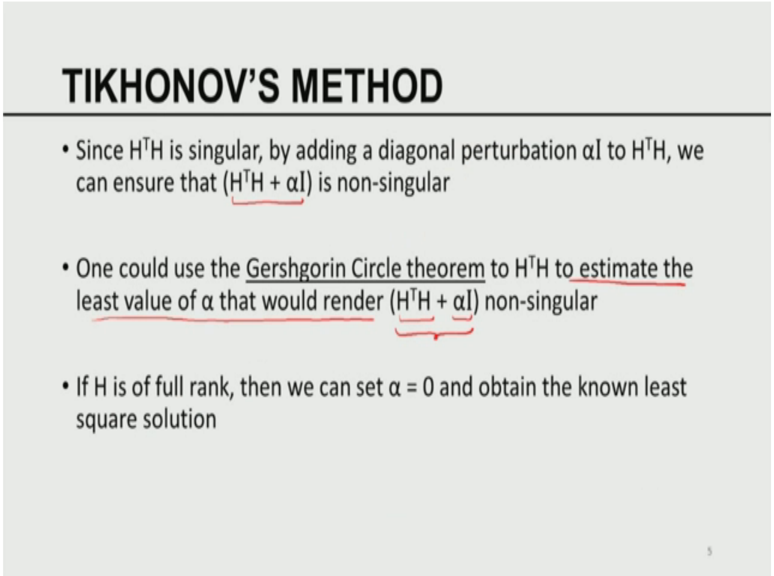
So, what is the Tikhonovs method? Tikhonov method is a modification of the method that we use in solving the under determined case. In the under determined case we want to, we seek to minimize the norm of H, but we require Z minus H to be a strong constraint and used Lagrangian multiplier. What did Tikhonov say? Tikhonov said you still consider factor that corresponds to the square of the norm of x, alpha is a parameter, Z minus H of x norm. So, this is square of the norm this is square of the norm, alpha is a kind of a penalty parameter. So, this is what is called a penalty function approach.

So, what is that we are looking for? I am looking for a solution f of, a solution that minimize f of x. The solution that minimize the f of x such that alpha times the square of the norm must be as small as possible and Z must be as close to H of x as possible, but not exactly 0. So, the addition of alpha x square term to the traditional sum of squared criterion helps to avoid the challenges resulting from rank deficiency.

Now, please understand in the over determined problem we f of x was essentially this is term for the over determined case, for the over determined case we essentially used the f of x to be this. Now, to that f of x I am adding this term by adding this term this is a kind of penalty term. So, we mix our (Refer Time: 09:00) and concoct a new objective function and our problem is to able to minimize this f of x. And let us talk about the impact of this after we solve the minimization problem. So, rewriting f of x as in equation 2 we can readily compute the gradient and equate the gradient to 0, if you equate the gradient to 0 I get the a solution for X LS, by setting this is to 0 I get the solution to be like this. So, you can see least square solution is a function of alpha H transpose H plus alpha I inverse H transpose Z, when you set alpha is equal to 0 it becomes a solution of the over determined case. So, you can think as of this generalization of concept of solution for the over determined case. So, this over determined case I considered only this term we added that. So, the alpha term is the one that is added to this.

Now, look at this now H transpose H is by itself is singular, but I am adding alpha kinds of identity matrices. So, this is called diagonal perturbation this change is only that diagonal elements of the matrix H transpose H. So, by adding a diagonal perturbation to a singular matrix I can make the whole matrix non singular if the matrix is non singular I can compute the inverse if you can compute the inverse I have the solution. So, what is this? We are not solving the original problem we are solving a modified problem the modification is obtained by adding a diagonal perturbation to the grammian H transpose H.

(Refer Slide Time: 10:47)

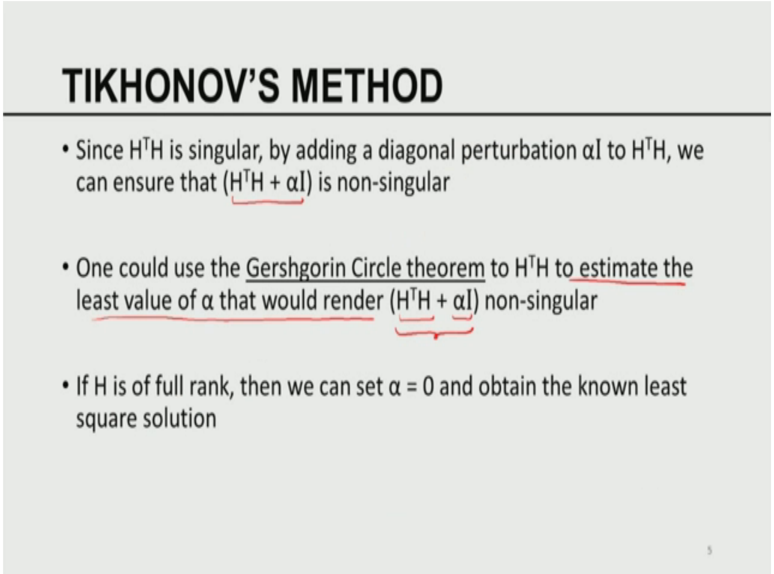


So, since H transpose H is singular by adding a diagonal perturbation alpha I to H transpose to H we ensure that H transpose H that plus alpha is non singular. When this is non singular I can compute the inverse, if I can compute the inverse I have an expression for the least square solution. So, the expression for the least square solution for this ill posed problem is given by is given by the equation 3 in slide 4, equation 3 in slide 4.

Now, you may ask a question how do I pick that alpha. In fact, I am going to give an intuitive fail, the alpha that you need to choose must be the smallest alpha that will make the matrix H transpose H non singular. This matrix by this matrix by itself is singular I am going to add perturbation I am going to require this whole matrix be non singular. So, I want to ask myself the question what is the least alpha that the I should use in order to render the resulting matrix to be non singular, such an alpha always exist it can be proven. So, this is a very nice generalization. So, when H is a full rank we simply set alpha is equal to 0 when H is rank deficient you pick the least alpha that will make this matrix non singular. So, once you have picked the least alpha that will make it non singular I can solve the linear least square problem I have a least square solution.

The existence of such alpha is guaranteed by as theorem in matrix theory called Gershgorin circle theorem. Using the Gershgorin circle theorem one can estimate the least value of alpha, one can estimate the least value of alpha that is needed to be rendered this is non singular such a thing exists it is a very simple result. So, what is that we trying to do? By formulating the problem as a penalty function problem we can even solve an ill posed problem nicely.

(Refer Slide Time: 13:13)



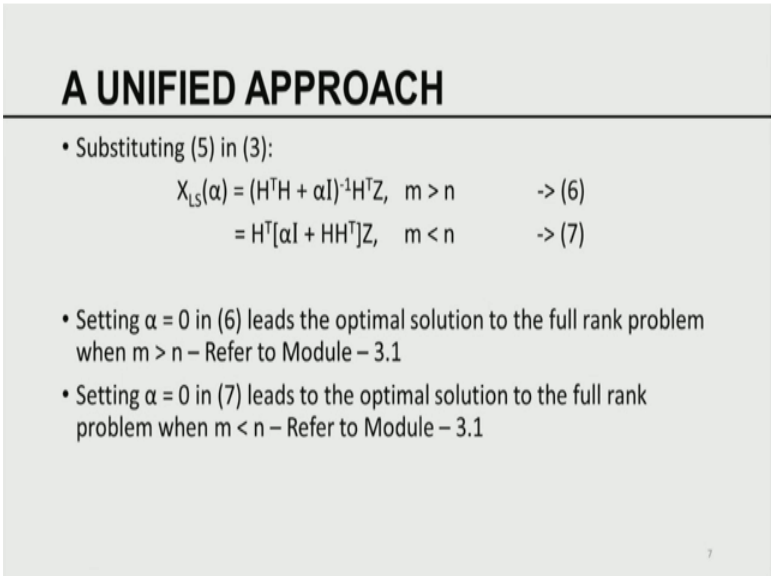
Now, I am going to talk about the use of matrix identity. I hope you all remember that we have talked about several different matrix identities when we dealt with the modular matrices. A very well known matrix identity takes this following shape by equation 4.

This identity is very well known in matrix theory, now I am going to start with the well known identity I am going to specialized A is equal to H, B is equal to I, D inverse is alpha I. Why? I would like to able to use this identity in the Tikhonov solution to see what is the relation between the under determined over determined case that is what our aim is. This identity if this substitution now becomes this using this identity, this left hand side is equal to the right hand. So, this left hand becomes this right hand becomes that the right hand can be simplified and that becomes this.

So, look at this now H transpose H plus alpha I inverse H transpose is equal to H transpose alpha I plus HH transpose inverse, these two problems these two matrices are equal that is the essence of equation 5. From here now I can do, I can see the lots of things if I said alpha is equal to 0 this becomes a solution for the over determined left hand side become a over determined system, the right hand side becomes a solution the under determined system.

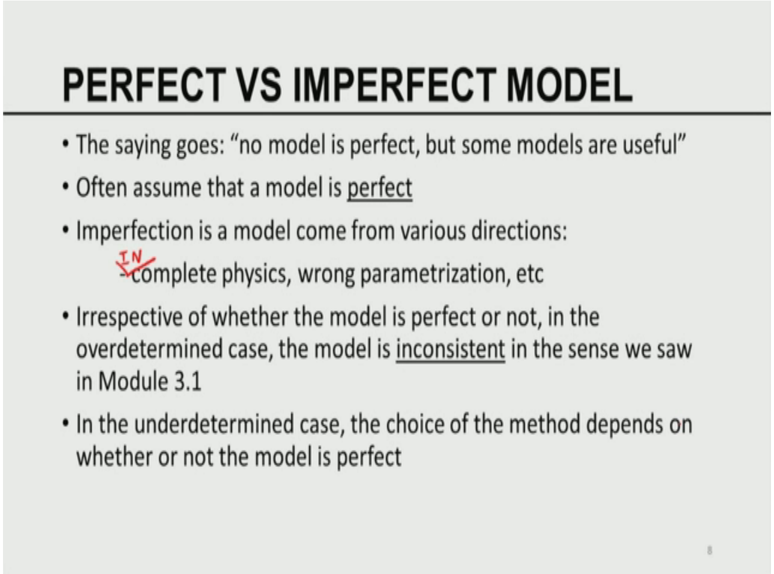
So, Tikhonov by introducing penalty factor alpha was able to use unify the solution for the over determine and under determine case by invoking to these very well known matrix identity. So, that is the beauty of the solution of Tikhonov. So, Tikhonov solution is important in two ways, one it helps to solve the ill posed problem another one it helps to unify in trying to define the relation between under determined over determined. So, you kill (Refer Time: 15:31) that is the beauty of the work by Tikhonov. Tikhonov has specialized in solving inverse problems of various types, he has written a marvelous book that deals with regularization methods for solving inverse problem. This is the one of the simplest of the method available regularization that is often used in solving linear least square ill posed problems.

(Refer Slide Time: 15:58)



So, I have already talked about the unified approach when alpha is 0 in 6 leads to optimal solution to the full rank problem when m is greater than n where alpha is 0 in 7 that leads to the optimal solution to the full rank problem when under determined case. So, all the previous solution can be obtained as special cases from the unified approach and hence then importance of Tikhonov of contributions.

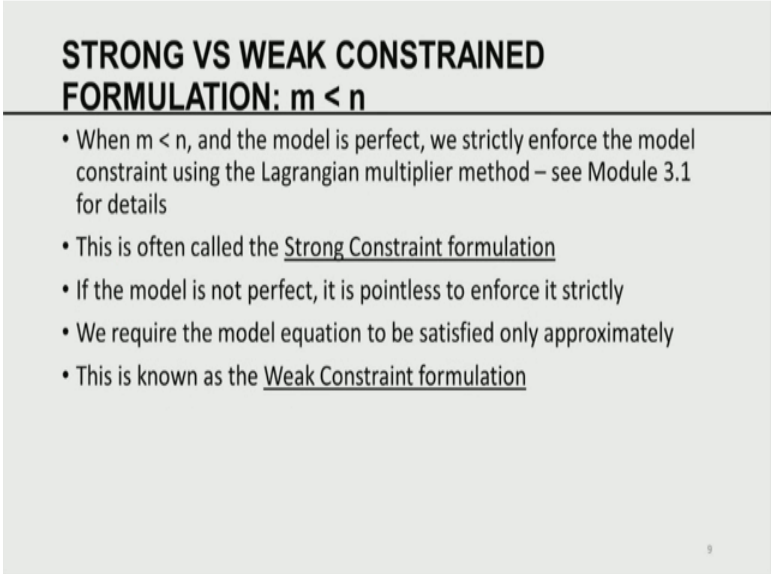
(Refer Slide Time: 16:25)



Now, I am going to talk about the role of model perfect verses imperfect under static constraint. So, models static, models can be perfect, model can be imperfect the saying goes no model is perfect, but some models are useful. Often one assume that a model is perfect, imperfection in a model comes from various directions. The imperfection from the model can come from incomplete physics not complete physics, sorry the imperfect model can be come from incomplete physics are wrong parameterization are unique combination that of are other reasons.

Irrespective of the model whether is perfect or not in the over determined case the model is always the problem is inconsistent, in the sense that we saw in the previous lecture over determined systems are always inconsistent. In the under determined case the choice of the method depends on whether the model is perfect or not. So, if the model is perfect we can format one way if the model is imperfect we can format other way. So, usual have a good feel for a good model is.

(Refer Slide Time: 18:02)

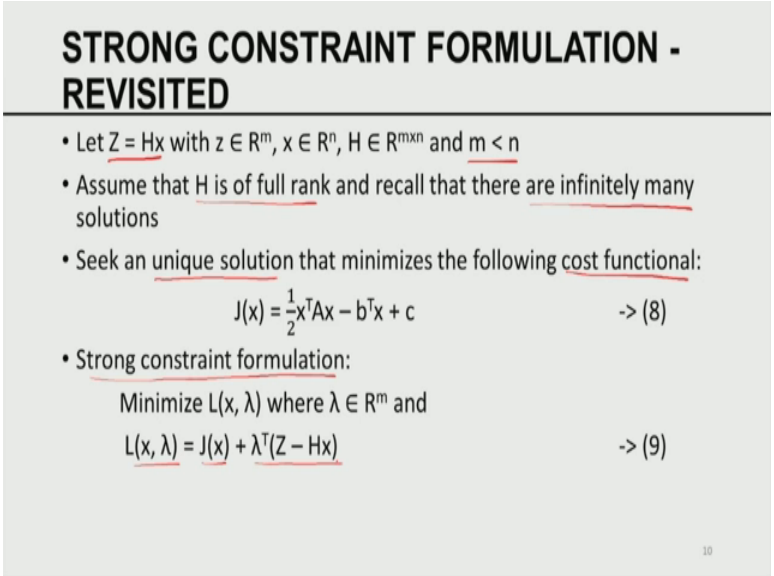


So, that brings us is the notion of different ways of formulating the solution of least square problems when the model is perfect when the model is imperfect, that gives rise to a new way of looking at it called strong constraint verses weak constraint formulation. When m is less than n the model and the model is perfect, so I am considering I am considering an under determined cases there are less observation than number of parameters and also a model is perfect and that the assumption we already made. We never question about the whereas it in the model until now, only now we are trying to ask the questions is the model perfect if the model is not perfect that is one way.

If the model is imperfect other way if the model is perfect I would like to able enforce the model equation strictly that gives raise to strong constraint the model constraint is strong using the Lagrangian multiplier. This version of using Lagrangian multiplier when you believe the model is perfect this called strong constraint formulation, we visualize strong constraint formulation in trying to bring uniqueness to the under determined system. If the model on other hand is not perfect it is pointless to enforce it strictly why we know the model is not perfect, why would you enforce something that you know you are not sure about.

So, in this case we still want to respect the model equation, but not strictly, but only approximately this ability to require the model equation to be satisfied not perfectly, but very closely is the concept of weak constraint formulation strong constraint weak constraint. Strong constraint is intimately related to, strong constraint formulation are formulated as a Lagrangian multiplier problem weak constraint problem are formulated as penalty function methods which we have already seen in the optimization module.

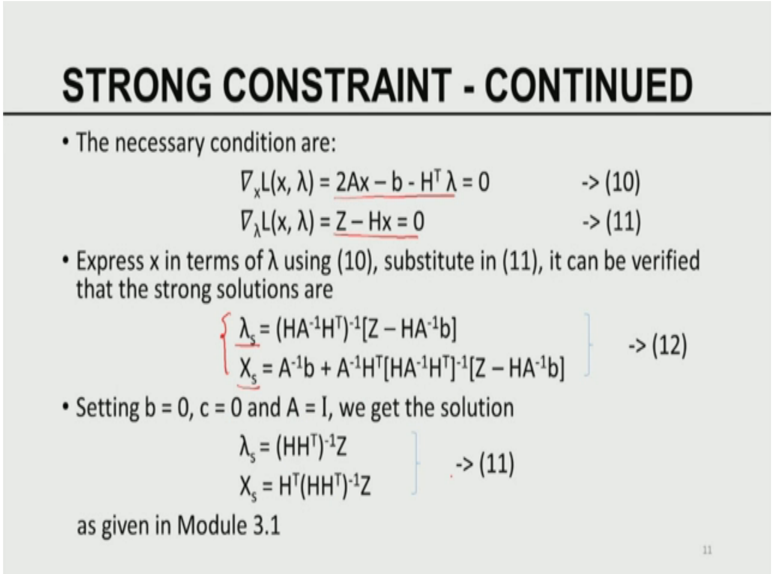
(Refer Slide Time: 20:05)



So, I am you going to quickly illustrate a version of the strong constraint formulation let Z is equal to H of x and considering the under determined case, assume H is a full rank, recall there are infinitely many solutions, we seek the unique solution that minimizes the following cost functional.

So, model is the constraint, among all the solutions of the model I want to find that solution that minimizes the cost function, the model has infinitely many solutions among the infinitely many solutions I am seeking a solution that minimizes this cost function. So, this cost functional generalization of the cost function that we utilize under, I already in the analysis of under determined system. In the under determined case what is it that we did? We would like to be a min find the x whose norm is minimum, but in here I would like to be a able to find an x such that it minimizes j of x which is general quadratic function. Strong constraint formulations for this a strong constraint formulation is to build a Lagrangian, the Lagrangian is j of x plus lambda transpose Z minus H of x.

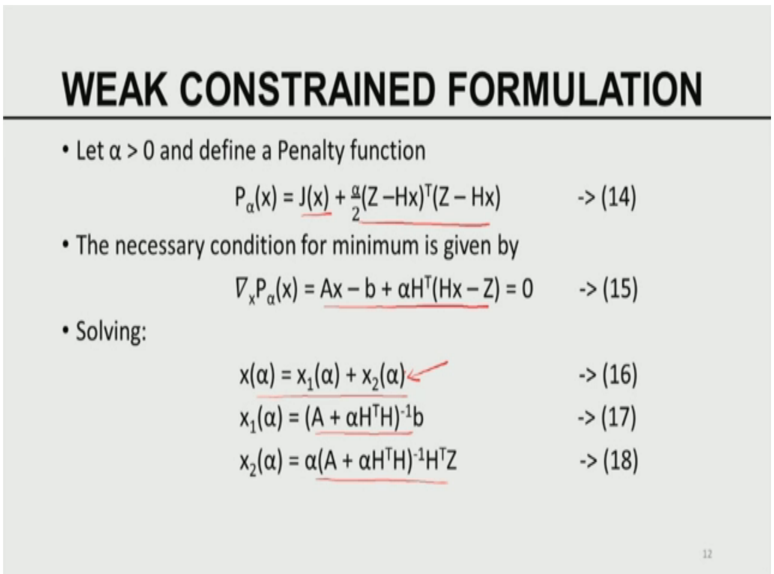
(Refer Slide Time: 21:30)



I am now, completing the derivative I am sorry under the gradient of L with respect to x and lambda, I get to two sets of equations, I get two sets, if equations I simultaneously solve these two systems I get the solution lambda s, s refers to strong solution, x of s refers to strong solution. The strong solution both lambda and x of s are given by the equation 12 which can be easily solved by, which can easily obtained by solving internal 11.

By setting b is equal to 0 c is equal to 0 and A is equal to I we get the well known solution for the under determined case which we have already seen. So, you can see this kind of generalization of what we are done in the under determined case.

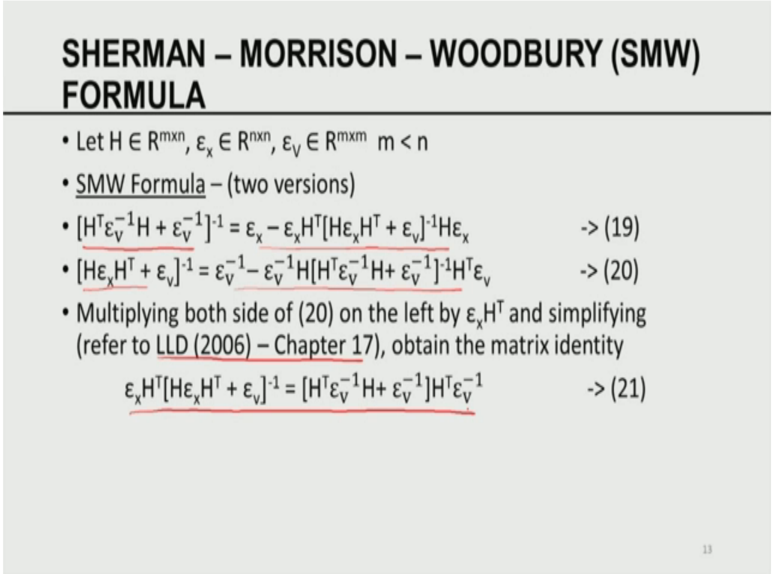
(Refer Slide Time: 22:21)



Now, let us consider the weak constraint formulation. In this case I am going to build a penalty function p of x that is equal to j of x plus alpha by 2 Z minus H of x transpose Z minus H of x. The necessary condition for a minimum is given by the gradient of p of alpha must be 0, the gradient of alpha for this is given by right hand side equate this to 0 and solving this you get the solution x 1, x is equal to x 1 alpha plus x 2 alpha, x 1 alpha has taken this form x 2 alpha has taken this form. So, the solution has two components both of which depends on alpha the sum of 17 and 18 is 16. So, 16 is the solution that minimizes the penalty function.

Earlier we talked about the relation between the weak solution and a strong solution. I know the form of weak solution, I know the form of the strong solution we also saw the weak solution tends to the strong solution as the penalty parameter alpha infinity now I am going to show that result in here.

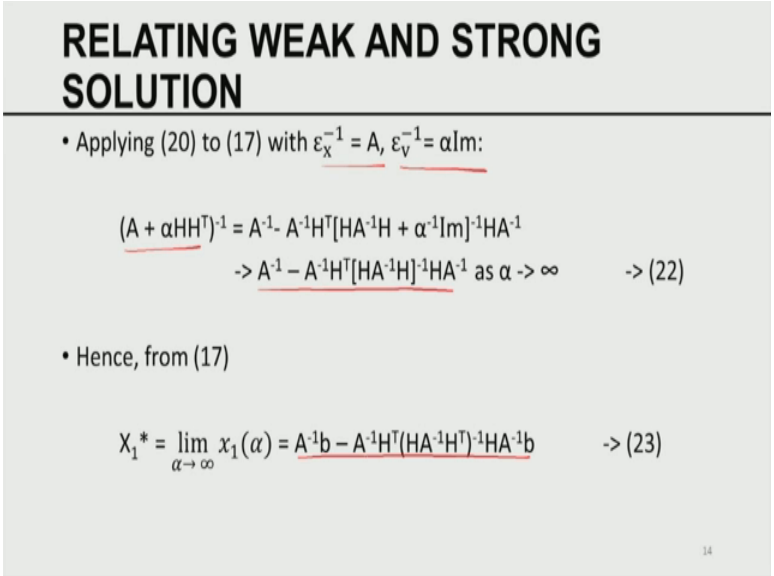
(Refer Slide Time: 23:31)



To this end we use the Sherman Morrison Woodbury formula, if I use the Sherman Morrison Woodbury formula you can see the relation on the left hand side this inverse applying the Sherman Morrison Woodbury formula becomes this, this inverse applying the Sherman Morrison Woodbury formula becomes this, we have already talked about the Sherman Morrison Woodbury formula. We also have given a proof of the Sherman Woodbury formula in the section of matrices.

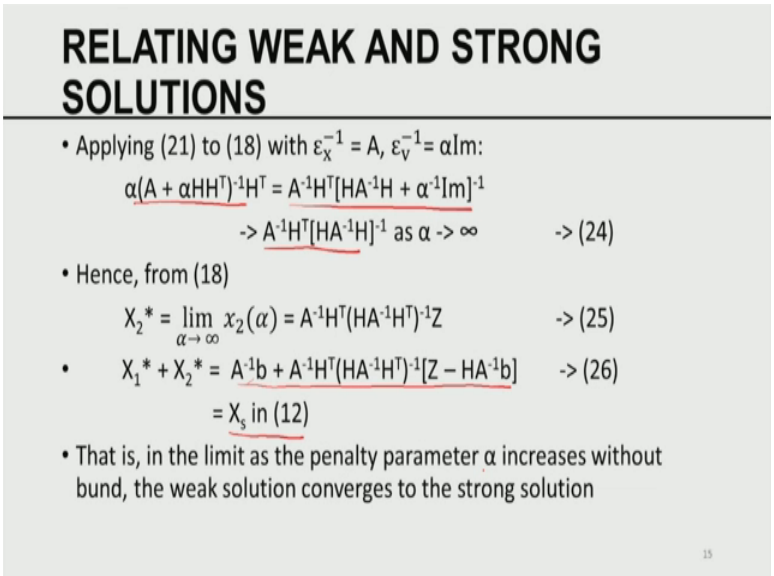
So, multiplying both sides of 20 on the left by epsilon x H transpose and simplifying in fact, I would like to refer the reader to details in chapter 17 of the book by Lewis Lakshmivarahan Dhall, our text book. We obtain the matrix identity which is given by this relation 21. I know there is lot of computational and checking is there, but I am sure I am checking is a homework problem, I am hiding on the all major concepts once you have the major concepts you can really follow these to be able to verify by computing various things.

(Refer Slide Time: 24:49)



Now, by setting these we can readily see this formula is given by this we can then we can see x 1 star is the limit of x 1 alpha which becomes this. So, look at this now when alpha goes to infinite the right hand side does not have independence on alpha. So, this is the limit of one of the components of the solution x 1 alpha in the limit.

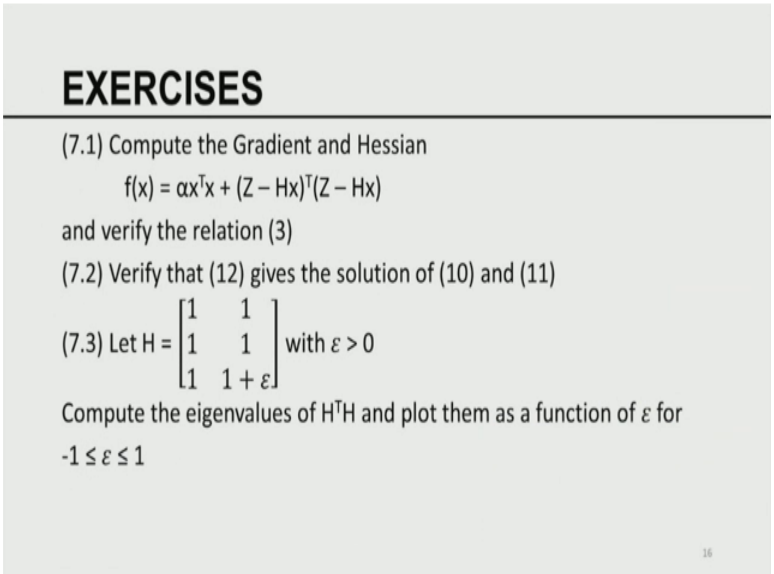
(Refer Slide Time: 25:20)



Likewise you can compute the solution x 2 alpha and using this identity again this matrix are identity if you have already described, using this matrix identity this becomes this. As alpha turns to infinity this becomes this, this is independent of alpha in the limit therefore, x 2 star is equal to limit of x 2 alpha which is given by that therefore, x 1 star plus x 2 star is given by this equation. If you simplify that, that becomes the strong solution X s in 12. So, we have demonstrated that the weak solution in the limit as alpha tends to infinite becomes a strong solution.

So, in other words as the penalty parameter alpha increases with that bound weak solution converges to the strong solution. So, where do I use the penalty function formulation? You believe in the model, but you also know the model is not perfect. So, whenever you have to use the model is the constraint use as a weak constraint. You know the model you assume the module is the perfect in that case you assume I am sorry you use the model as the strong constraint, so which method do you use that depends on how strongly you believe on the goodness to the model, on the goodness to the model. Therefore, we have now described how to take care of many many very many special cases.

(Refer Slide Time: 26:50)

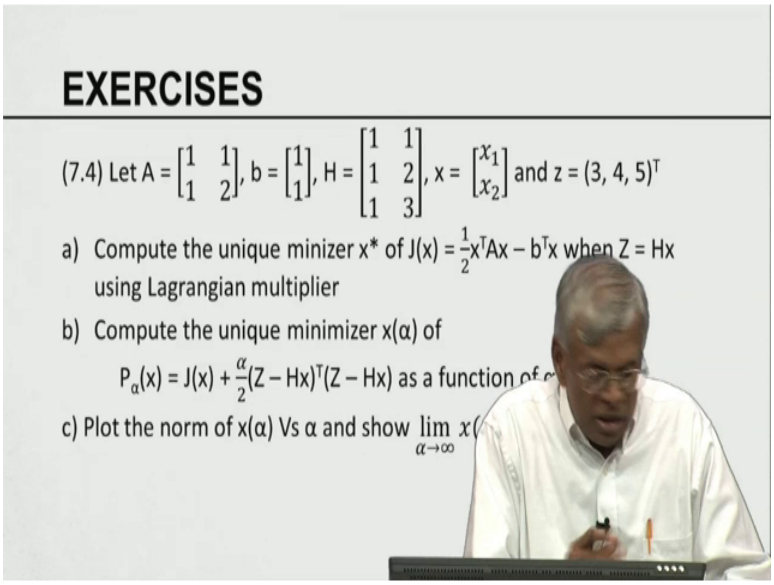


So, I would like to summarize observing the following. The linear static deterministic inverse problem can be broadly divided into well posed and ill posed. Well posed problems are very straight forward, ill posed problems in principle can cause headache. There are multiple ways of solving ill posed problem, one way would be to use Tikhonov of regularization as we have talked about, the second way would be to use the singular valued composition we have not done that we will wait until that is done to revisit this issue. We also then incorporated the notion of perfect and imperfect models. So, you can now consider quite of right situation, models is perfect, model is static, model is perfect imperfect, model is deterministic, model is linear. So, you have linearity, perfectness or imperfectness, you have well posed or ill posed, over determine or under determine. So, you can consider a quite a variety of formulation of the even simple linear least square problems that brings about the beauty that underly the notion of simple linear least square methods in the context of linear inverse problems.

I would like to very strongly recommend that you all work out the exercise these are simple extensions of the methods that we have solved. I am particularly referring to example 9.3 in here I am considering a matrix whose columns are 1 1 1, 1 1 1 plus epsilon. When epsilon is 0 the two columns are linearly depended, when epsilon is not 0 of mathematically it is linearly independent therefore, the eigenvalues of H transpose of H are functions of epsilon. So, what is that I would likely to do, H is given by this epsilon greater than 0 epsilon small compute to the H transpose H, compute the eigeenvalue of H transpose of H. Plot the variation of eigenvalue using a MATLAB as epsilon varies from minus 1 to plus 1.

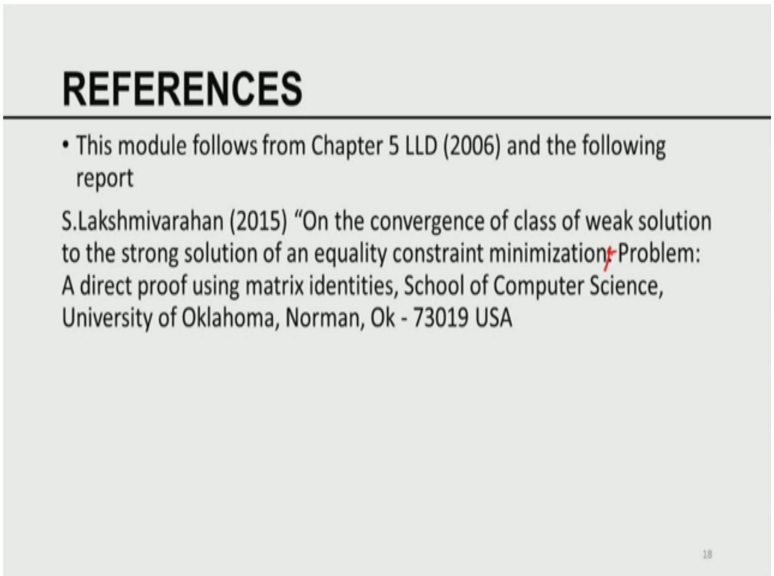
So, you can see what is the impact of the rank. So, when the epsilon is 0 the rank, this is the rank deficient when epsilon is not, it is not, it is strictly rank 2. So, but when epsilon is small even though mathematically it is rank 2, it can still cause problems and how it creates the problems one can understand by computing the eigeenvalues. Once you compute the eigenvalues you remember we can compute the condition number, therefore, by computing the condition number for H transpose H we can infer how ill posed or how well posed the problems are. So, that is the measured by which we can compute the degree of ill posedness. So, this problems 7.3 is an important, is an important exercise.

(Refer Slide Time: 30:14)



There are few other problems which are routine, which follow the directions of the development.

(Refer Slide Time: 30:22)



And this modules follows from chapter 5 and the following the report. We recently compute report in 2014 on the convergence of the class of weak solution to the strong solution of an equality constraint minimization problem, equality constraint minimization problem I am sorry that is, this must be we must remove the point there. A direct proof using matrix identities is a technical report from School of Computer Science, University of Oklahoma.

With this I think we have provided a broad over view of the richness of the linear least square problems static version thereof.

Thank you.