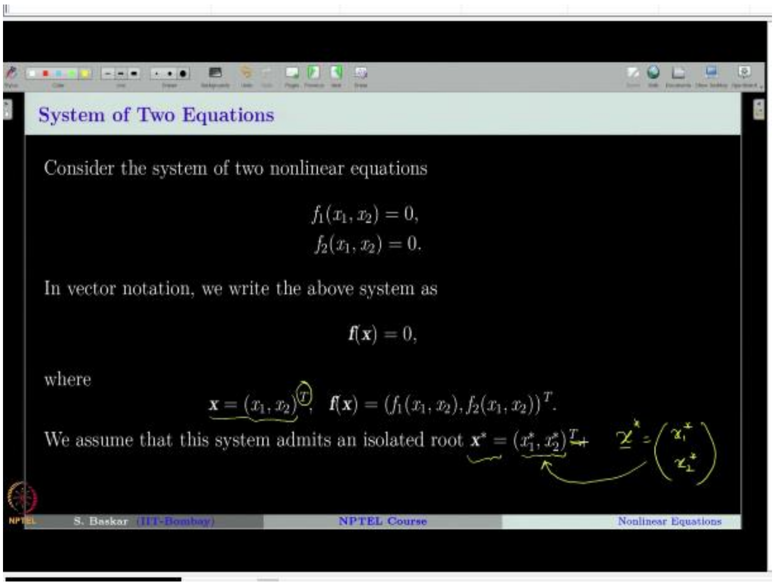
**Numerical Analysis**
**Prof. Dr. S. Baskar**
**Department of Mathematics**
**Indian Institute of Technology-Bombay**

**Lecture-36**
**Nonlinear Equations: System of Nonlinear Equations**

Hi, we are discussing numerical methods for capturing isolated roots for non-linear equations. In today's class we will extend our studies to system of equations. We will restrict our discussion only to system of two equations. Once we understand this, extending the idea to any number of equations is straightforward.

**(Refer Slide Time: 00:44)**

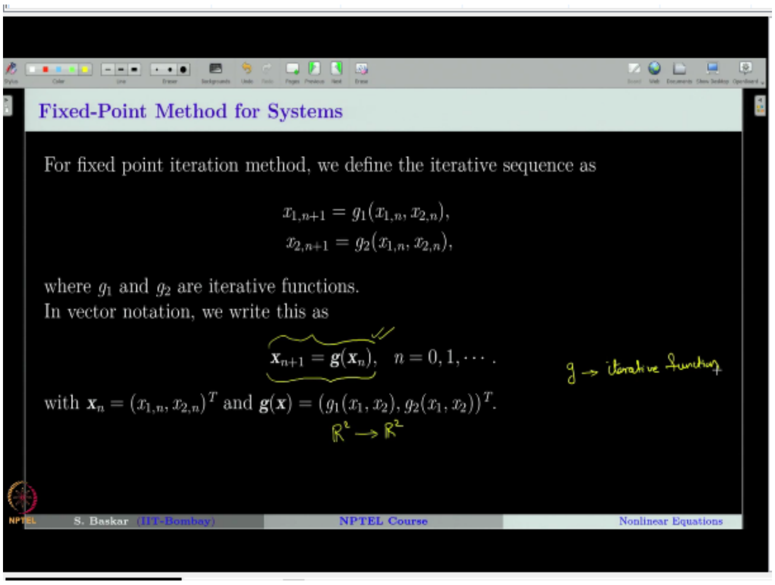



So, with that in mind let us consider two non-linear equations given by $f_1(x_1, x_2) = 0$ and $f_2(x_1, x_2) = 0$, where $f_1$ and $f_2$ are functions from $\mathbb{R}^2 \to \mathbb{R}$. We can write this system in the vector notation as $f(x) = 0$, **0** is vector and **f** is given by $(f_1, f_2)$ and **x** is the vector in $\mathbb{R}^2$ and **f** is a function from $\mathbb{R}^2 \to \mathbb{R}^2$. Now we are interested in finding an isolated root of this non-linear system.

Let us call this isolated root that we are interested in as the vector $\boldsymbol{x}^*$ and it is given by $(x_1^*, x_2^*)^T$. Remember, right from our chapter on numerical linear algebra we always considered vectors as column vectors. That is why whenever I write a vector, I write it in the row wise and then put a $T$ there.

It means it is written in the column wise, I am doing this just to save the space. If I write it in the column wise it may take more space. So, you should understand the vector for instance this as $(x_1^*, x_2^*)^T$. This is what we mean by writing like this and putting a transpose here.

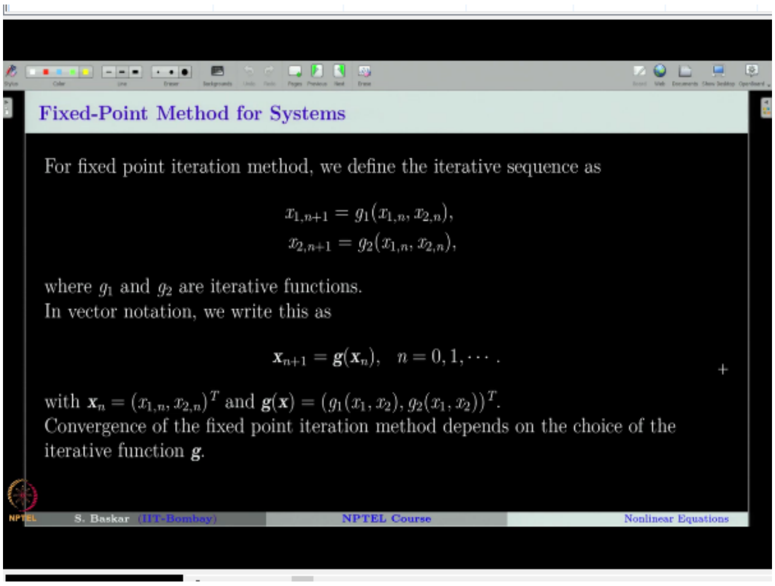
**(Refer Slide Time: 02:47)**




The idea is to first set up a fixed point iteration sequence. If you recall in our last two lectures we have introduced this general framework of iterative methods with one initial guess in the form of a fixed point iterative method. So, we will apply that idea component wise to each equation in our system and thereby we have two equations. In the vector notation this fixed point iteration method can be written like this.

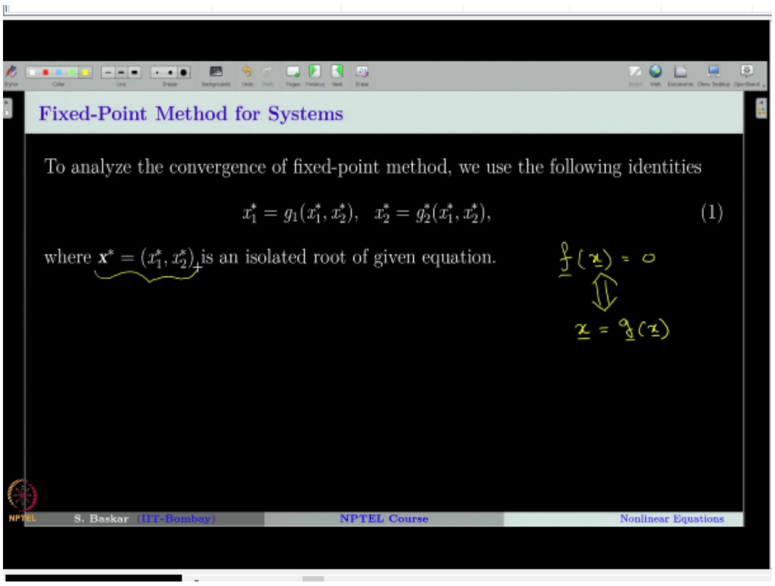
Where $\mathbf{g}$ is the vector valued function from $\mathbb{R}^2 \to \mathbb{R}^2$ and $\boldsymbol{x_n}$ is a vector in $\mathbb{R}^2$ and this expression generates the iterative sequence. We will provide the value of $\boldsymbol{x}_0$ as the initial guess and then we will generate this sequence using this expression for $\boldsymbol{x}_1, \boldsymbol{x}_2, \boldsymbol{x}_3$ and so on. Remember the function $\mathbf{g}$ is called the iterative function and this has to be specified as the part of our method itself.

**(Refer Slide Time: 04:14)**




All these ideas are familiar to us. Therefore, I will not explain them once again.

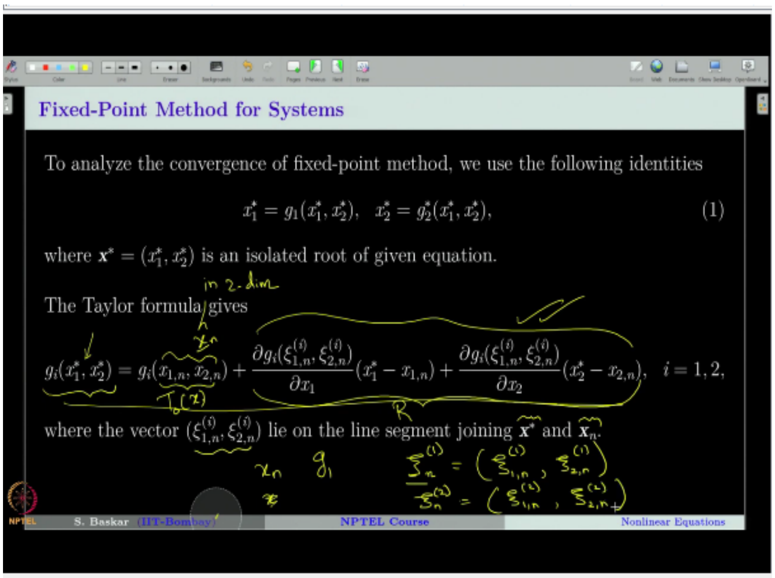
**(Refer Slide Time: 04:20)**




Now to analyze the convergence of this fixed point iteration method, if you recall we have certain hypothesis proved in our last class when we worked with single non-linear equation. The same idea will also go with system of equations. Recall that you have this system $\mathbf{f}(\boldsymbol{x}) = 0$ you will write it in an equivalent form as $\boldsymbol{x} = \mathbf{g}(\boldsymbol{x})$. That is the main idea of the fixed point method.

So, you will choose this iteration function in such a way that isolated root of the non-linear system $\mathbf{f}(\boldsymbol{x}) = 0$ is also an isolate fixed point of the iterative function $\mathbf{g}$. Therefore, we will denote the isolated fixed point of $\mathbf{g}$ as $\boldsymbol{x}^*$.

**(Refer Slide Time: 05:20)**



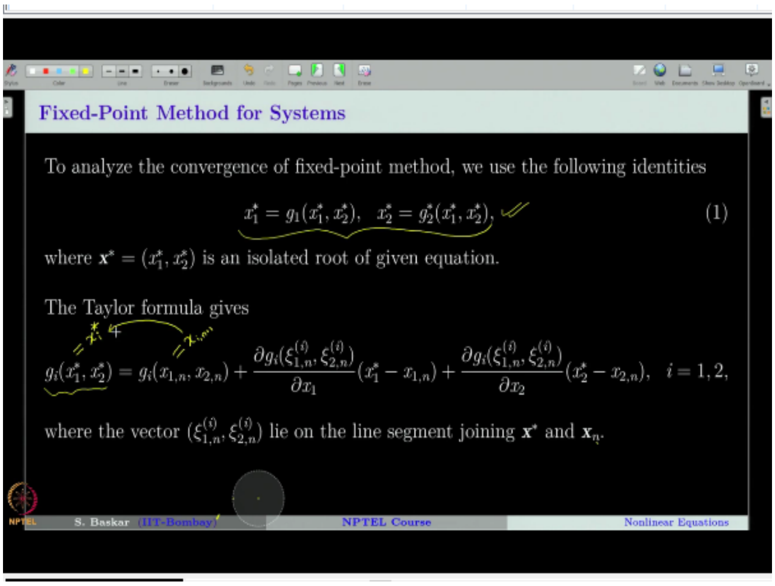

Now we will use the Taylor approximation to write each component of the vector **g** around an iteration vector $\boldsymbol{x}_n$ with Taylor polynomial being degree 0. Then the reminder term is given like this. Remember this is the Taylor formula, in two dimensions; you should go back and recall how we write Taylor's formula in 2D. We have only introduced it for one variable. You can similarly write the Taylor formula for two variable functions.

Here I am just writing the function **g** at the fixed point and that is taken around the point $\boldsymbol{x}_n$ and we have only taken the Taylor polynomial of degree 0 and therefore the remainder term is written like this. Here you will get two $\boldsymbol{\xi}_n$'s; one for each component that is for $g_1$ when you expand you will have a $\boldsymbol{\xi}^{(1)}$ and since we are expanding this **g** around the point $\boldsymbol{x}_n$ therefore we will also use the notation $\boldsymbol{\xi}_n^{(1)}$ for it.

And this is a vector and it is given by $\xi_{1,n}$ and this is for the first component of the vector **g** and $\xi_{2,n}$ and this lies on the line segment joining $\boldsymbol{x}^*$ and $\boldsymbol{x}_n$. Remember you are writing **g** at $\boldsymbol{x}^*$ and using the Taylor formula at $a = \boldsymbol{x}_n$. That is what you are doing and the Taylor formula is written with polynomial of degree 0 that is Taylor polynomial of degree 0 and you have the remainder term.

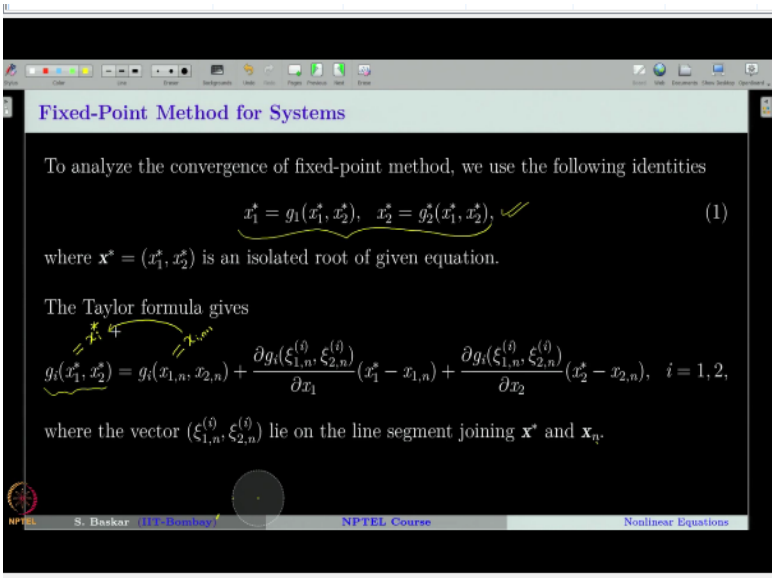
So, you have to do it component wise. We have two components for **g**, $g_1$ will give one such expression and $g_2$ similarly will give another expression like this with $\boldsymbol{\xi}_n^{(2)} = (\xi_{1,n}^{(2)}, \xi_{2,n}^{(2)})$. This is the notation that we will use.

**(Refer Slide Time: 07:56)**




Now you can see that this is equal to $x_1^*$, this is what we have even here; **g** is chosen in such a way that $\boldsymbol{x}^*$ is a fixed point. Therefore, this is equal to $x_i$, for *i* = 1, $g_1$ of this is $x_1^*$ and $g_2$ of this is equal to $x_2^*$. That is what we have written here. Similarly using the iteration formula, you can write this as $x_{i,n+1}$.

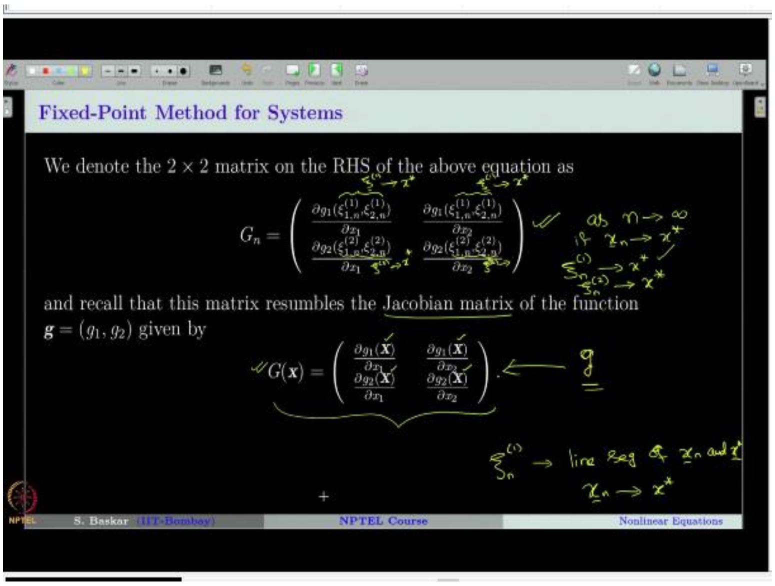
**(Refer Slide Time: 08:43)**




And then, I will take this to the other side to write $x_i^*$ which is on the left hand side $x_{i,n+1}$, which I have taken from the right hand side to the left hand side and that is going to be equal to the remainder term. That is already there on the right hand side. So, that is what we get here and this is obtained component wise. That is *i* = 1 and 2. Just to write these two equations in the matrix form we get this equation.

I am just putting $i$ = 1 and that gives me the first row of this system and put $i$ = 2 that gives me the second row of the system. So, now we got this. Now we will give a notation for this matrix. Remember this is nothing but $\boldsymbol{x}^* - \boldsymbol{x}_{n+1}$ and this I will give a notation $G_n$ therefore this will be $G_n(\boldsymbol{x}^* - \boldsymbol{x}_{n+1})$. So, this is what we can write it in the vector notation.

**(Refer Slide Time: 10:06)**

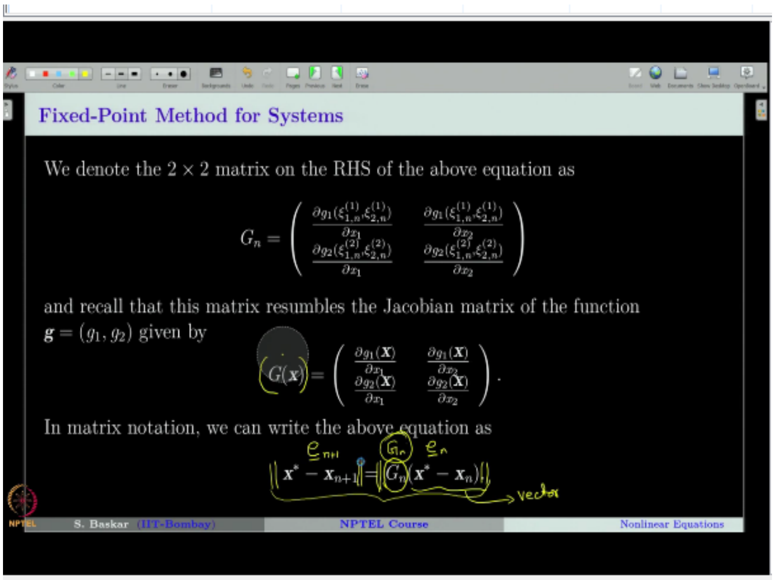



For that I will define the matrix $G_n$ like this. Now what is this $G_n$? If you recall the Jacobian of the function $g$. Remember $g$ is a vector valued function, therefore you can write the Jacobian matrix of $g$ and it is given like this. Now this is not exactly the Jacobian matrix because for having the Jacobian matrix you have $x$ applied in all the elements of this matrix, whereas here we are applying $\xi^{(1)}$ for the first row.

This is also $\xi^{(1)}$ and this is actually $\xi^{(2)}$. So, this $G_n$ actually resembles the Jacobian matrix of G, but it is not exactly G, but one interesting thing that we can observe here is as $n \to \infty$ your sequence $\boldsymbol{x}_n$ converges to say $x^*$ then you can see that $\xi_n^{(1)}$ will also converges to $x^*$ and $\xi_2^{(1)}$ will also converges to $x^*$. Why it is so? You can see that $\xi_n^{(1)}$ and $\xi_n^{(2)}$ are lying on the line segment of $\boldsymbol{x}_n$ and $x^*$.

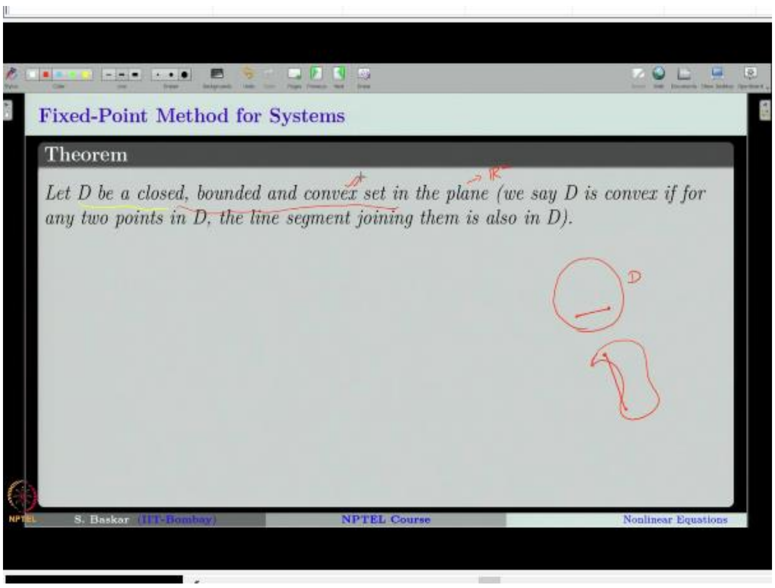
So, that is what we have seen when we were expanding $g$ in the Taylor formula. Therefore as $\boldsymbol{x}_n$ tends to $x^*$ by the Sandwich theorem you can see that $\xi_n$ will also converges to $x^*$ Therefore, in the limiting case everybody will have $x^*$ as the argument. That is all the elements of this matrix will be evaluated at $x^*$ in the limiting case. So, in the limiting case this $G_n$ will be precisely the Jacobian of the matrix g evaluated at the fixed point $x^*$ of this function $g$.

**(Refer Slide Time: 12:40)**




And as I told in the previous slide, we can write our equation in the vector notation like this. Remember this is the error vector at the $n$ + 1th iteration that is equal to $G_n\boldsymbol{e}_n$. Remember this is a matrix and here you see this is a matrix and this is a vector. Therefore, the right hand side is the product of matrix and vector and that is actually a vector. Now to see the convergence of the iterative method we have to take the norm of this equation on both sides. And then we have to see how this matrix behaves.

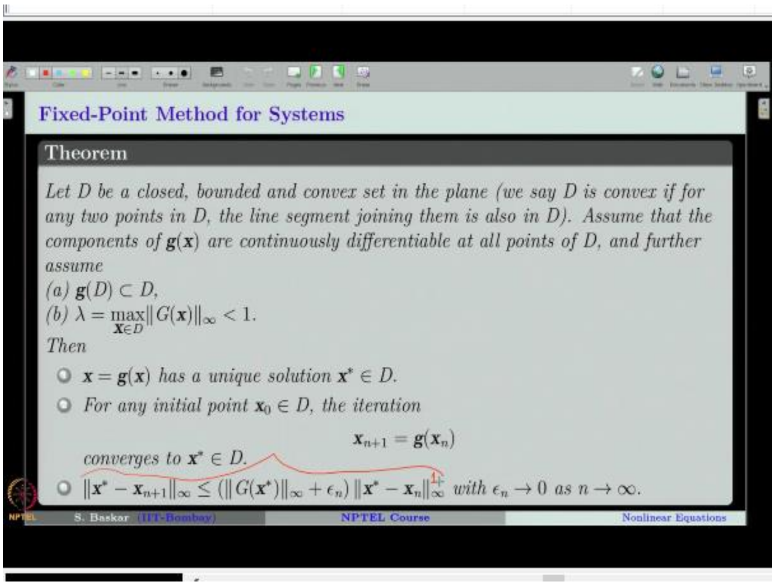
**(Refer Slide Time: 13:35)**




Let us state the convergence theorem of the fixed point method for a system and we will omit the proof of this theorem. The theorem says that let $D$ be a closed bounded and convex set in $\mathbb{R}^2$. So, if you recall what is mean by a convex set? By convex set we mean that the set $D$ should be in such a way that you take any two points in the set $D$ and draw the line segment joining these two points.

That line segment should entirely lie in that set and this should happen for any pair of points. So, suppose if you take a set like this, this is not convex because if you take two points like this and draw the line segment that line segment will not entirely lie in the set $D$. So, this is used in the proof of this theorem. That is why we are assuming.

**(Refer Slide Time: 14:43)**




Anyway, we will omit the proof for this course. The second assumption is that we will assume all the components of the vector $g$ are as continuously differentiable functions at all the points in the domain $D$ and further we assume that the iteration function $g$ is a self map. If you recall we now know why we are assuming all these conditions on $g$. At least in the case of single non-linear equation we have studied all these properties in full details.

And then if you take the maximum of the matrix norm remember this is the subordinate matrix norm and we are just restricting it to the $l_\infty$ norm that does not matter you can use any norm. For easy handling in the proof it is better to have the $l_\infty$ norm. That is why we have taken this and if you take the maximum of the $l_\infty$ norm of $G$ then that should be less than 1.

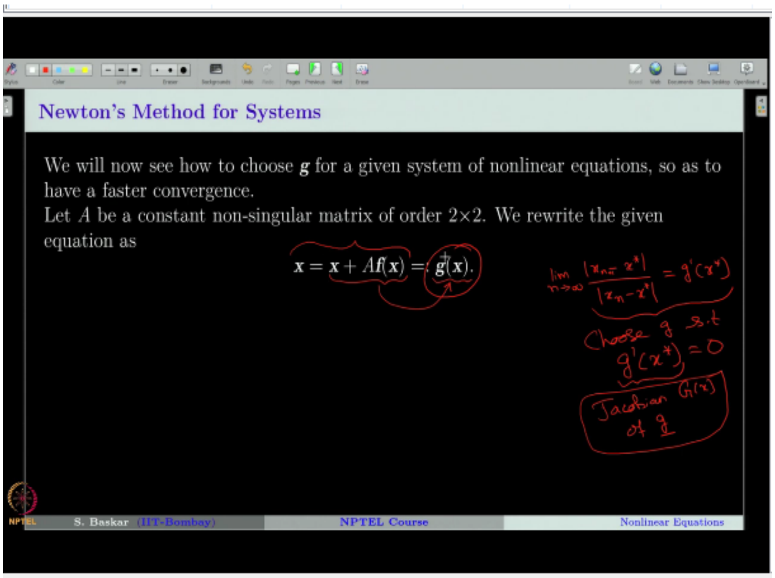
Remember this is the contraction condition. Then the theorem says that $g$ has a unique fixed point in the domain $D$. Remember the existence comes from the continuity of the function $g$ and the self map and the uniqueness comes from the assumption that $g$ is a contraction map. Then for any initial condition $\boldsymbol{x}_0 \in D$. That is if you start your iteration by taking $\boldsymbol{x}_0$ in the domain $D$

then the iteration sequence $\boldsymbol{x}_n$ will be well defined and all the terms of the iterative sequence will lie in the domain *D*.

And further $\boldsymbol{x}_n$ will converge to the fixed point $x^*$ in *D* and further we also have this estimate which tells us the order of convergence of this fixed point iteration method is actually 1. That is it has a linear order of convergence. So, you just keep your knowledge that you had in the case of single non-linear equation and just see this it is only the vector notation but idea is the same as we have studied in the case of single non-linear equation.

However, proof will use some multivariable calculus ideas therefore we will omit the proof of this theorem. Now the question is can we improve the order of convergence by suitably choosing our iterative function $\boldsymbol{g}$? Again if you recall we have done such an exercise in the case of single non-linear equation again the idea goes very much similar to that.

**(Refer Slide Time: 17:52)**

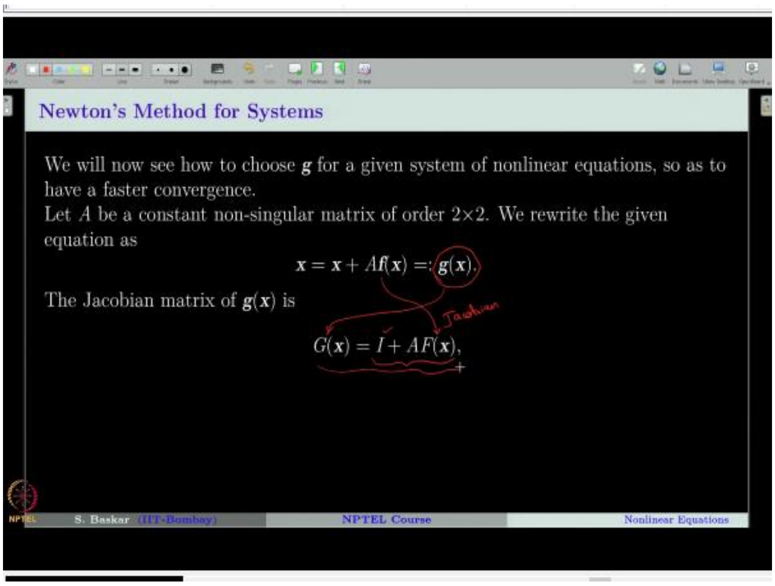



Let us see we will now see how to choose $\boldsymbol{g}$ for a given system of non-linear equation so as to faster the convergence. For that we will now take $\boldsymbol{g}$ that is our iteration function as $\boldsymbol{x} + A\boldsymbol{f}(\boldsymbol{x})$ and look for the fixed point of this function. So, this is your $\boldsymbol{g}(\boldsymbol{x})$ and you are looking for the fixed point of $\boldsymbol{g}(\boldsymbol{x})$. That is you are looking for the root of this non-linear equation, where *A* is a matrix that needs to be determined in order to faster the convergence.

If you recall, how can we make a faster convergence in the single equation. Let me quickly recall that we had $\lim_{n\to\infty} \frac{|x_{n+1}-x^*|}{|x_n-x^*|} = g'(x^*)$. That is what we have obtained. Now in order to have higher order convergence here we need to choose g such that $g'(x^*) = 0$.

That is the idea the same idea here instead of $g'$ in the single equation. Now we will have the Jacobian of the vector valued function $\boldsymbol{g}$. So, that is the only difference that we will have in our present case. Let us see we will find the Jacobian of this iterative function and see how it looks like.

**(Refer Slide Time: 19:49)**

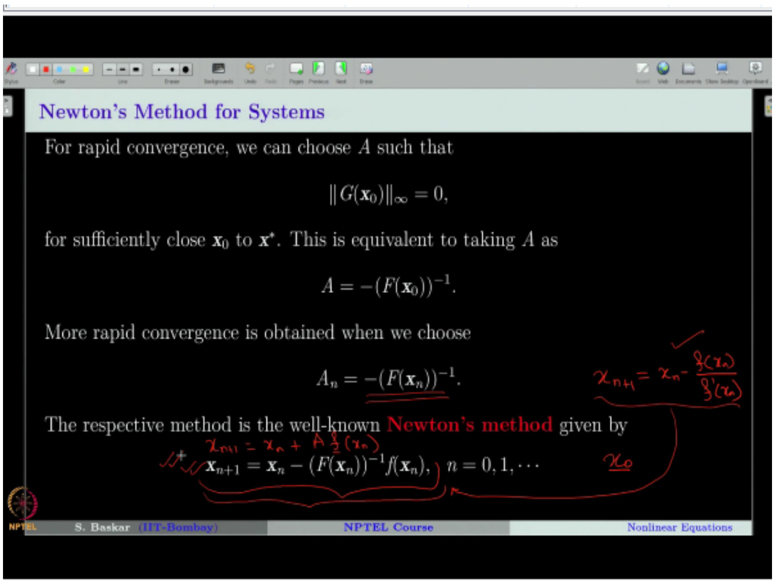



You can see that the Jacobian of this function is $\boldsymbol{g}(\boldsymbol{x})$ and it is given by the $n \times n$ identity matrix $I + AF(\boldsymbol{x})$, where $F$ is the Jacobian of the function $\boldsymbol{f}(\boldsymbol{x})$. So, that is how the Jacobian will look like.

**(Refer Slide Time: 20:17)**

And if you go back recall that the Jacobian of the function $\boldsymbol{f}(\boldsymbol{x})$ is given by this matrix. Now how to get the faster convergence. Precisely we have to look for that iterative function $\boldsymbol{g}$ such that this is equal to 0. That is the idea which is similar to what we had in the single equation case.

**(Refer Slide Time: 20:48)**




So, therefore we will look for the iteration function **g** such that norm of $||G_n||_\infty = 0$. Let us try to achieve that for our initial guess $\boldsymbol{x}_0$. Now how you will get, remember $\boldsymbol{g}(\boldsymbol{x_0}) = AF(\boldsymbol{x}_0)$. Now we want norm of this to be equal to 0. That should be the case when the matrix itself is the 0 matrix. Remember this is one of the conditions of the matrix norm.

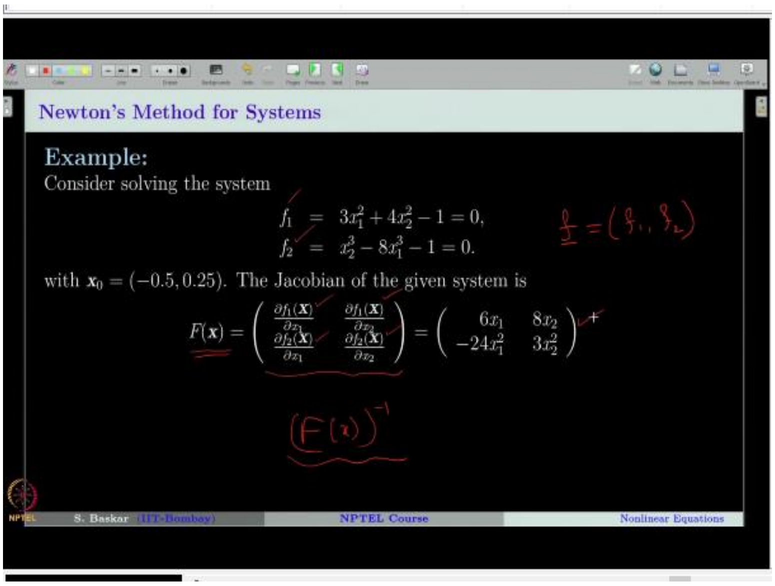
So, I am just using that in order to have the matrix norm. Remember this is the subordinate matrix norm. So, in order to have the matrix norm equal to 0, I should have my matrix itself as 0. So, if this matrix has to be 0 it means my $A = -\left(F(\boldsymbol{x}_0)\right)^{-1}$. So, that is how I will get this expression to be 0. So, therefore in order to have this which will actually make my fixed point iteration sequence to have a higher order of convergence.

In order to have this, I need to choose my $A = -\left(F(\boldsymbol{x}_0)\right)^{-1}$ and this I will do for each iteration. That is, I will propose my $A$ just to have the notation $A_n$ here because I am going to change my $A$ in every iteration. Therefore, I will have $A_n$ which is defined as $-\left(F(\boldsymbol{x}_n)\right)^{-1}$ where $F$ is the Jacobian of the vector valued function $\boldsymbol{f}$. Therefore, my actual, iteration formula that is fixed point iteration formula is given like this.

Remember we have taken it as $x_{n+1} = x_n + A\boldsymbol{f}(x_n)$ and now I have taken $A$ as this expression, therefore this is my iteration formula. Given $\boldsymbol{x}_0$ now I can generate my iteration formula for $n$ = 0 which gives me $\boldsymbol{x}_1$, $n$ = 1 that gives me $\boldsymbol{x}_2$ and so on. You can just see that this is very much similar to what we defined as Newton-Raphson method in the case of single non-linear equation.

In the case of single non-linear equation, the Newton-Raphson method looks like this $x_n - \frac{f(x_n)}{f'(x_n)}$ and this seems to be a natural extension of this formula to a system of non-linear equation and that is why this method is called Newton's method.

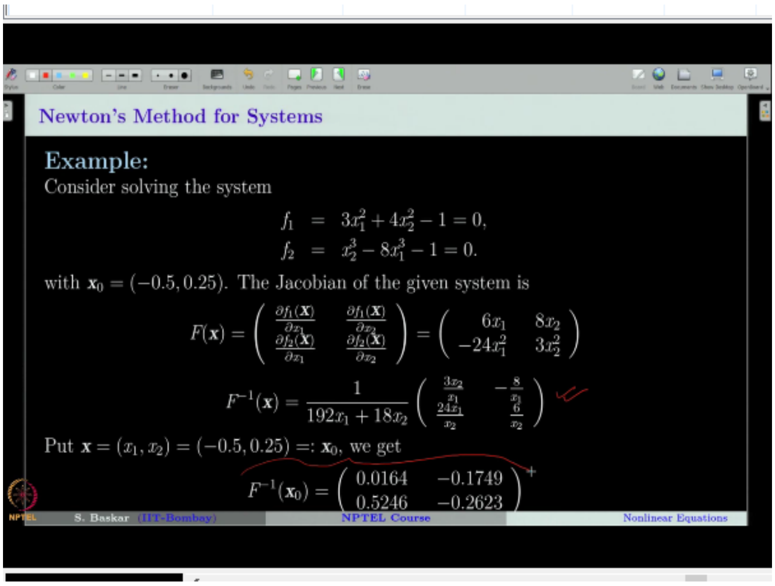
**(Refer Slide Time: 24:01)**




Let us do an example. Consider this system of equations and we are interested in finding an isolated root of this system. For that we will use the Newton's method with $\boldsymbol{x}_0$ given by this vector. As a first step we have to compute $\left(F(\boldsymbol{x})\right)^{-1}$, because that is involved in our formula. So, first we will find the Jacobian of the matrix $F$. Remember the vector $\boldsymbol{f}$ is nothing but $(f_1, f_2)$.

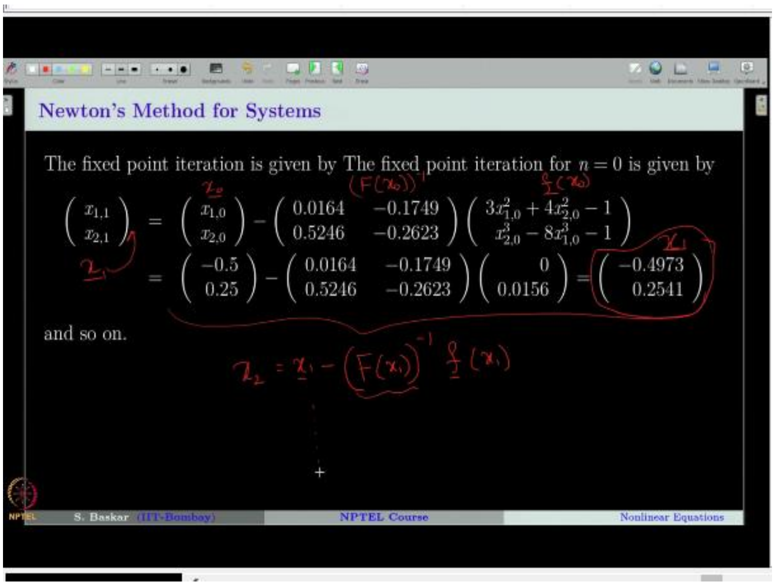
Therefore, the Jacobian of this function $\boldsymbol{f}$ is given by this formula and now $f_1$ is given like this and $f_2$ is given like this. From there you can compute each of the elements of this matrix and they are given like this. So, we found $F$.

**(Refer Slide Time: 25:38)**




Now we have to find the inverse of that matrix and that is given like this. Once you find the inverse of the Jacobian matrix $F$ you can now go to set up your formula for the Newton's iteration. Now for each $\boldsymbol{x}_n$ therefore you have to find this matrix, for $\boldsymbol{x}_0$ the matrix is given like this.

**(Refer Slide Time: 25:32)**

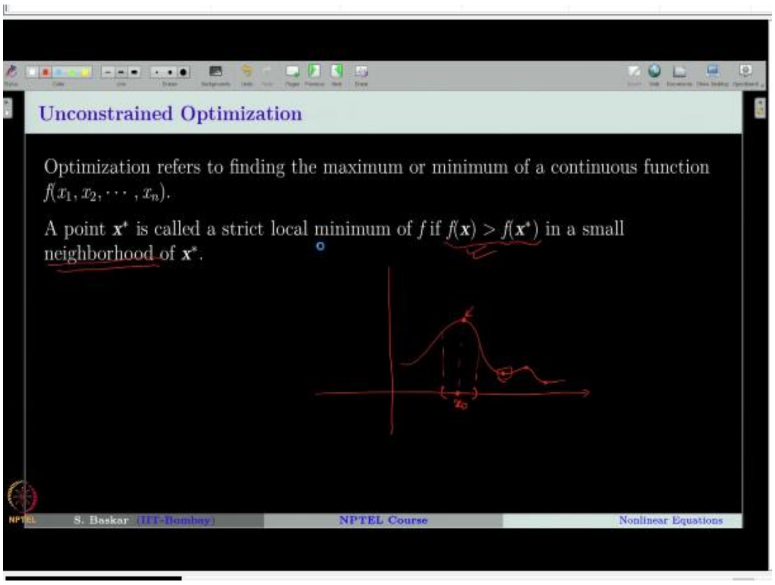



Therefore, to compute $\boldsymbol{x}_1$ which is component wise written like this and that is given by $\boldsymbol{x}_0$ which is this into $(F(\boldsymbol{x}_0))^{-1}$. So, remember this is obtained from the inverse of the Jacobian matrix times $F(\boldsymbol{x}_0)$. So, that is what is this and therefore the first iteration formula is given like this. Now once you have this idea you can compute your $\boldsymbol{x}_1$ and that is given like this.

Now once you have $\boldsymbol{x}_1$ you then have to find $F(\boldsymbol{x}_1)$. Once you find this matrix then you have to find the inverse of that. Once you find the inverse of that then you can set the expression for

$\boldsymbol{x}_2$ that is given by $\boldsymbol{x}_1$ minus this inverse matrix into the vector $\boldsymbol{f}(\boldsymbol{x}_1)$. Like that you can keep on going as long as you want to go. So, this is how we can implement the Newton's method for a system of two equations.

**(Refer Slide Time: 26:47)**

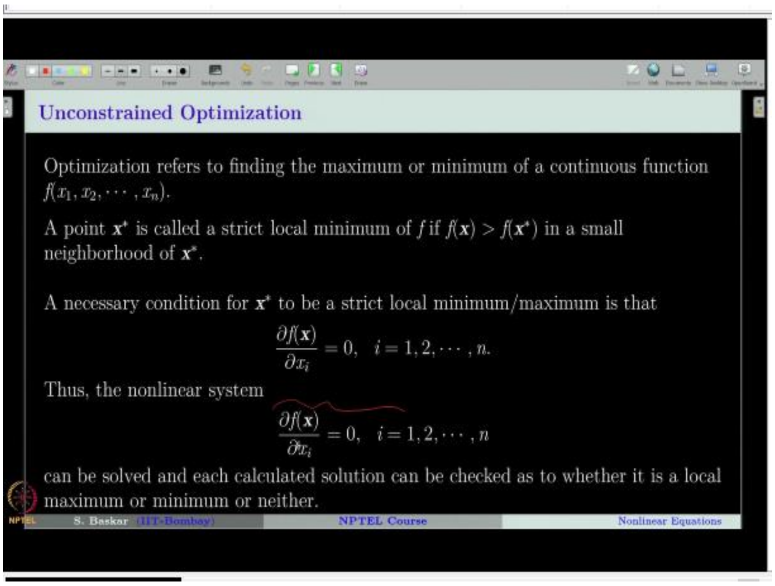



Let us take a small application of this Newton's method. Here we would like to find an extremum for the given real valued function $f$ defined on $\mathbb{R}^n$. So, this is what is called the unconstrained optimization problem. We will assume that the function $f$ is continuous function and recall that a point $\boldsymbol{x}^*$ is called a strict local minimum. Similarly, you can also define for local maximum.

A point $\boldsymbol{x}^*$ is called a strict local minimum of $f$ if $f(\boldsymbol{x}) > f(\boldsymbol{x}^*)$. Suppose you have the function like this, I am just showing it in the case of function in $\mathbb{R}$ then this is a local maximum and this is the local minimum and this is local maximum and so on because you just have to have this condition in a small neighbourhood of $\boldsymbol{x}^*$.

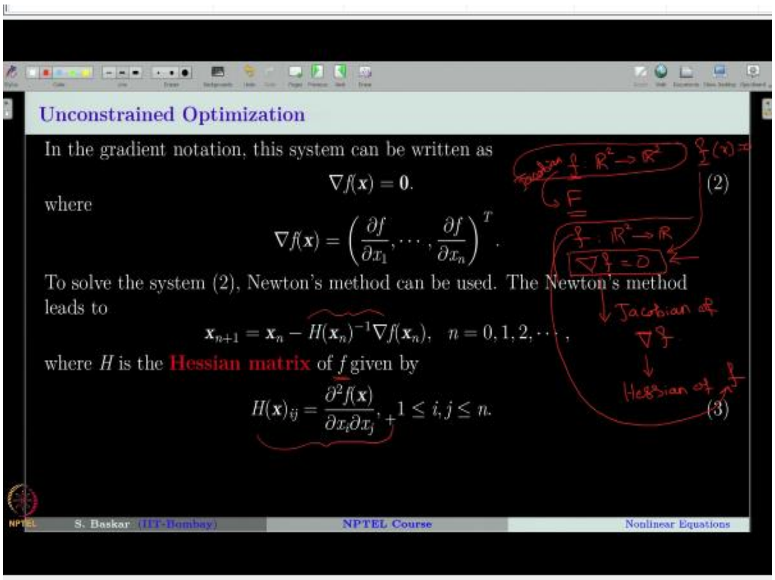
Suppose this is your $\boldsymbol{x}^*$ then you can see that in a small neighbourhood of this $\boldsymbol{x}^*$ this is the maximum. So, therefore this is a local maximum and similarly you can see that this is a local minimum and so on. So, similarly you can also define what is mean by local maximum.

**(Refer Slide Time: 28:18)**




Now from the calculus course you can see that a necessary condition for $\boldsymbol{x}^*$ to be a strict local minimum or maximum is that this partial derivative should be 0, thus we have to set a non-linear system of equation like this. So, that an isolated root of this system will be a local maximum or minimum of the given function.

**(Refer Slide Time: 28:52)**



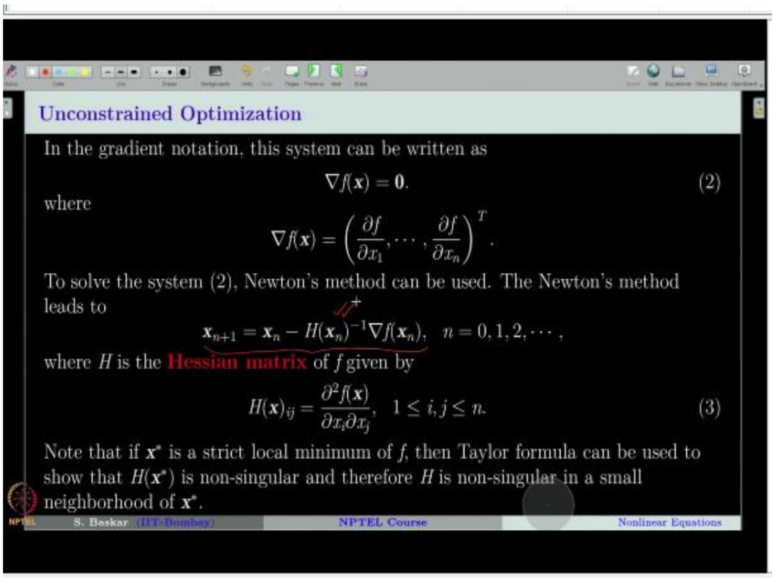

So, let me write it in the vector notation. If you see this gives you the gradient vector. Therefore, I am writing my non-linear system as $\nabla f(\boldsymbol{x}) = \mathbf{0}$ vector, where the gradient is given like this. This is precisely what I am writing this system in the vector notation. Now we have to find an isolated root of this system. For that we can use the Newton's method to obtain an approximation to it.

And the Newton's method is given by this formula where $H$ is the Hessian matrix for the function $f$. If you recall in our first discussion, we had a $\boldsymbol{f}$ that is the function from $\mathbb{R}^2 \to \mathbb{R}^2$ and we had Jacobian of this and its inverse is used in the Newton-Raphson method. Now what we are doing is there is a slight confusion in the notation that we are still using the notation $f$ now which is a function from $\mathbb{R}^2 \to \mathbb{R}$.

And we want to find a local maximum or minimum for this function. For that what we are doing is we are setting up this system $\nabla f = 0$. So, that is what in the unconstrained optimization we are doing. In unconstrained optimization we are given a real valued function defined on $\mathbb{R}^2$, in general it can be $\mathbb{R}^n$ and then to find the local maximum or minimum of this function we have to set this non-linear equation.

Therefore, you have to apply the Newton-Raphson method for this equation just like how you applied the Newton-Raphson method for this equation $f(x) = 0$ in our previous discussion. Similarly, you have to now apply this Newton-Raphson method for this equation and therefore you have to find the Jacobian of $\nabla f$. Jacobian of $\nabla f$ is nothing but the Hessian matrix of your given function $f$. So, that is why we are having this and that is given by this formula.

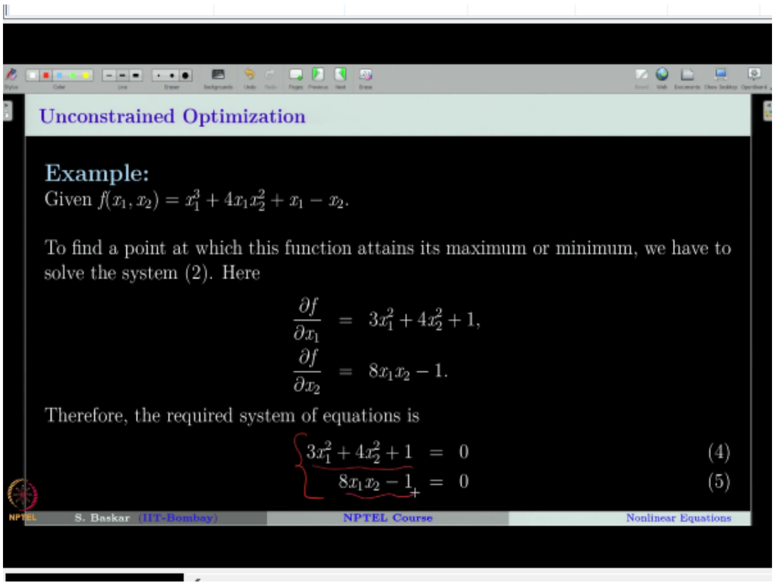
**(Refer Slide Time: 31:42)**




So, once you have the Hessian matrix then you know how to set up the Newton's method iterative sequence and it is given like this. Just follow the Newton's method that we have introduced in our previous slide. Note that if $\boldsymbol{x}^*$ is a stick to local minimum of $f$ then Taylor formula can be used to show that the Hessian matrix is non-singular.

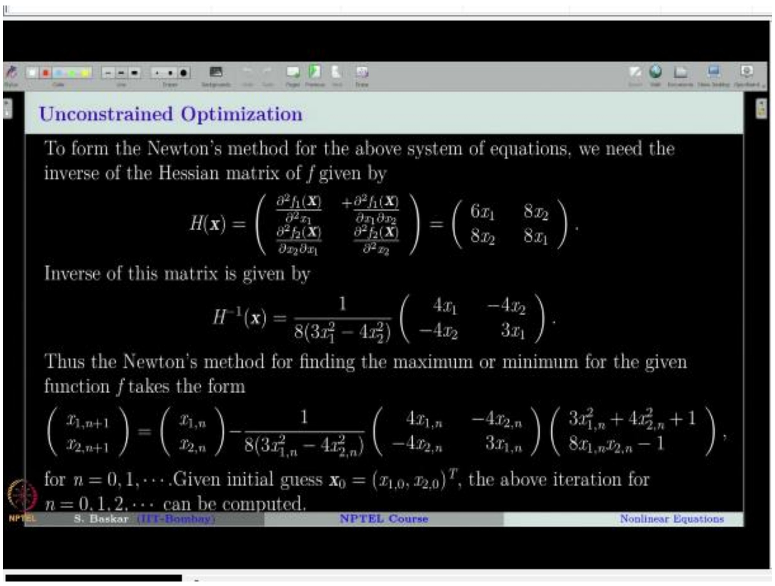
Therefore, you can always invert this matrix and therefore $H$ is non singular in a small neighbourhood of $\boldsymbol{x}^*$ and therefore you can always invert this matrix.

**(Refer Slide Time: 32:22)**




Again, let us quickly see an example. Now we are given this function, this is a real valued function defined on $\mathbb{R}^2$. Now we want to find a local maximum or minimum we will call it as local extremum of this function. For that first you have to set up your non-linear equation. For that you have to find the gradient of this function $f$ and it is given like this. Therefore, you have to set up the Newton's iterative formula for this system. So, you have to find the Jacobian of this vector.

**(Refer Slide Time: 33:05)**




And that will be the version of your given function $f$ and that is given by this matrix. Once you have this you can find the inverse of this matrix and that is given like this. Once you have the inverse of the Hessian matrix you can set up the Newton's method and that is given like this. Now once you have this then you can start with the initial guess and you can go on computing the Newton's iterative sequence.

And if that sequence converges it will converge to a local extremum of the given function $f(x_1, x_2)$ equal to this expression. So, that is the idea. With this our discussion on non-linear equation is complete. I thank you for your attention.