Observer Design for Nonlinear systems

Welcome back. In the previous class, I talked about the observer design, and we basically saw how to design the Luenberger observer for a linear time-invariant system. In this class, I am going to extend the same concept for the non-linear system. Again, our assumption is that the model of the nonlinear system is perfectly known to us. And then I am going to design the Luenberger-like observer for the nonlinear system. So, the purpose of the discussion is to design an observer for nonlinear systems.

And what kind of objective am I going to tackle in this lecture? I am going to address the state estimation challenge in the nonlinear system. Why? Because most of the time, full state measurement is not available. After that sensor deployment, we are measuring the state or output using only the sensor. So, that is costly in several ways and is also impractical in nature.

So, what are we basically going to do? We are going to design an observer to reduce the number of practical sensors. After that, we have uncertainty several times. So, this challenge I am going to tackle in the next class, but I will give you an overview of why this analysis is not valid when uncertainty comes into the picture. Obviously, whenever we are dealing with a non-linear system, we have to take extra care in terms of the theory. And due to that reason, we have to explore something new beyond the observability condition that is called uniform observability.

So, it is possible to show that if some assumption like uniform observability is satisfied, then I can design an observer for the non-linear system. Obviously, for a linear time-invariant system, this condition is Lipschitz continuity. Now, whenever non-linear system, then we have to actually incorporated the condition of Lipschitz continuity. So, we are going to see their significance, and after that, I have told you that whenever we are designing an observer or observer-based control, we have to prove the stability in terms of the error. So, since system is non-linear in nature and due to that reason, we are going to explore the Lyapunov theory.

And using this lecture, we are making some kind of bridge between theory and practice. So, I am going to tell you about systematic design; after that, I will also discuss stability analysis and implementation considerations. Our design methodology, obviously a step-by-step observer construction, is very crucial for gain matrix selection and handling of uncertainty, which is also important. The stability condition basically consists of the Lipschitz constant, Lyapunov-based convergence means, some kind of terminology that will come into the picture in terms of the Lipschitz and Lyapunov functions, and obviously, the robustness consideration. And obviously, this methodology somehow gives the room that even if you do not have the measurement of all states, you can design some kind of feedback that is based on all state information.

And at that time, convergence was very, very crucial. So, we are also going to analyze the convergence of the nonlinear observer. So, let us start this lecture. So, in order to design an observer, after obtaining the mathematical model, we have to express the mathematical model in some specific form. Because I have to check some kind of condition such that, if I have information about the input and information about $y(t)$, using this information, I can know the information of all initial conditions or, equivalently, the information of all states in finite time.

That is the basic motivation behind observability. So, what have we basically done here? First, after obtaining the mathematical model of the system, during a state-space representation, I represented the system in this way. You can see that this side contains all nonlinear parts, and that it depends on the control and the output only. Okay, and here I am assuming that I have $n$number of states, $m$number of controls, p number of outputs, and the pair $(A, C)$is observable. Okay, again I am going to explain the meaning of this, and ψ is locally Lipschitz.

So, this condition is very, very crucial. Okay, so before designing the nonlinear observer, first I have to check it, and $x(t)$and $u(t)$should be defined for all time. So, you can see that I have just information about $y(t)$and $u(t)$, but we are assuming that $x(t)$ should also be defined for all time, and for that reason, I have to provide some kind of guarantee in mathematical terms so that I can satisfy the last assumption. Now, the first assumption that is very crucial is the observability assumption, and I have already told you the physical interpretation of observability: that a state can be uniquely determined from the output and the input, provided this algebraic condition is satisfied.

It means that this is referred to as the Kalman rank condition. So, you can take the matrix $C$from the output and A, and check whether this matrix has full rank or not. If the answer is yes, then you are easily able to proceed towards the observer design. You can see here that in the previous lecture, when I was talking about the Luenberger observer, at that time we had just assumed this kind of assumption, but now I am going to incorporate one more assumption that is called uniformly observable. Okay, so what is the meaning of uniformly observable? Suppose you have some system; okay, now if.

.. You have fixed your input, and if you have given some kind of initial condition such that $x_1 = 0$, corresponding to that initial condition you will get $y_1(t)$. And what you can do is take the exact same system and apply the control. Now, you can change the initial condition; then obviously your output is going to change. So, the meaning of uniformly observable is that if your initial conditions are different, then these outputs are different for all $t \geq 0$. It means that if I know the output, I can understand the behavior of the initial condition or the behavior of the state regardless of the input; that is the meaning of this. And this condition is very essential whenever you are designing any kind of nonlinear observer.

So, here I am going to talk only about a very, very special class of nonlinear systems. But suppose that whenever you are going to talk about a hybrid system or a quantum system, you are trying to design the observer at that time. So, at that time you have to actually make this kind of assumption. Uniform observability ensures that the system state can always be reconstructed from the output, even in the presence of arbitrary input. So, input does not matter here.

And I have already told you that this property is very important. Now, in several places I am writing, I have used this terminology, that is locally Lipschitz, several times in this lecture. So, what is the mathematical definition and what is the physical interpretation of that? I am just going to revise it again. Suppose that if I have some function, which is a function of $u$and $y$in this particular application, you can see that ψ is a function of the control as well as the output. I am saying that this particular function is called locally Lipschitz.

In its argument, the arguments are $u$and y. So, for every $u > 0$and $y > 0$, if there exists some kind of neighborhood and a constant l, such that I am able to formulate an inequality of the form
$\| \psi(u_1, y_1) - \psi(u_2, y_2) \| \leq l \ \| (u_1, y_1) - (u_2, y_2) \|$,
and this inequality is based on a norm, then this is called locally Lipschitz. So, any norm can be taken here, and after that, you have to show that there exists some constant $l$locally. Okay, so basically what is the meaning of this condition? This ensures that whatever nonlinear function we are dealing with does not have a slope that is infinite. Okay, so due to that reason, this condition gives us the guarantee of existence and uniqueness of the solution.

Okay, so it means that your system is somehow well posed in nature. Due to this reason, this condition is very crucial whenever you are designing the observer for the nonlinear system. So, let us summarize what kind of conditions are required, or what kind of conditions we have to check, before proceeding toward the observer design. So, the first step, obviously after doing the mathematical modeling, is to express the system as
$\dot{x} = Ax + \psi(u, y)$,
and after that,

$$y = Cx$$

in this form. So, if we are lucky enough to express the system in this way, then I have to check the uniform observability.

So, obviously, the algebraic condition is observability, and after that, by taking several inputs and initial conditions, we have to show that if the initial condition is not exactly the same, then the output is also not exactly the same. Here, we also have to assume that both the input and the states are bounded because of the non-linear system. So, we do not know the nature and due to that reason, these kind of assumptions give us the flexibility whenever we are talking about the guarantee of error convergence or well-behavedness of the whole coupled system because during observer design, we have already seen in the previous class that extra dynamics comes into picture. Obviously, regarding the Lipschitz property, we have already seen its significance because, by using this property, I can actually prevent the finite time blow-up. So, if you look at the literature on observer design for non-linear systems, it is possible to show that several times, if your observer design is not proper, then that observer design will somehow destabilize your original system.

It means that in finite time you can be able to move to infinity. So, that condition can be prevented by using this kind of assumption. And continuity is obviously important. And after that initial condition. One more important point is here.

Since in order to prove the stability of whole system. So, our system with some kind of observer design. So, we are relying on the Lyapunov analysis because during control design the separation principle is not obvious, and for that reason, we have to restrict our set of initial conditions. This is also a valid assumption because most of the non-linear models are only valid in some specific space. So, I am assuming that the initial condition should belong to some kind of compact set.

So, once we confirm that all assumptions are verified, we will proceed to the observer design. And you can see here that the observer design for the nonlinear system is exactly like that of the LTI system. So, what I have done is copy the dynamics, and here I have actually changed x to $\hat{x}$, and after that, I am going to add the correction term. And again, the philosophy is exactly the same: as $y$tends towards $C\hat{x}$, at that time this term is not going to contribute anything, and at that time $x$is actually tending towards $\hat{x}$. And what is $\hat{x}$? That is the estimated state vector.

So, this program is going to run inside a computer. This is our original plan. So, it is possible to show that by designing H, which is nothing but the gain of the observer, this system composed of components running inside a computer is exactly the same as the original plant. And in this way, whatever state is actually coming from here, after a very long time, you can assume that both are equal; initially, you are not able to do this, because $y$and $C\hat{x}$are not equal. So, this term is not equal to zero, and for that reason, $\hat{x}$is different from x.

Now, I have to minimize the difference between $x$and $\hat{x}$, and for that reason, I have to define some kind of error between these two terms, and I have to force this error to be equal to 0. So, like in the previous class, it is possible to show that this problem is exactly the same as placing the poles of $A - HC$towards the left half-plane. So, I have to design Hsuch that all eigenvalues of $A - HC$are strictly negative. So, in this way, I am able to design the observer. Now, since this term is known, due to that reason, if you calculate the error dynamics, then the error dynamics $\dot{\tilde{x}}$ is nothing but $\dot{\tilde{x}} = (A - HC)\tilde{x}$, and after that, $\tilde{x}$ comes into the picture.

And due to that reason, I am able to comment like this. So, now what I am going to do is take a very simple example. So, the matrix $A$is like this. Here, I am assuming that this is a second-order system. So, whenever I am fixing
$C = [1 \quad 0]$,

it means that I am assuming that I have information about x_1.

Why? Because
$r_1 Cx = y$,
which gives
$y = x_1$.

Now, using that information, I am going to estimate the information of the second state, which is x_2. That is our goal. So, what step do I have to follow? First, I have to check the rank of the observability matrix. So, the Kalman matrix, which is also called the observability matrix, is constructed, and after that, if you calculate the rank of this matrix, it turns out to be equal to 2.

So, I have a second-order system, and the rank of this matrix is 2. What does it mean? It means that I am able to proceed towards the observer design. Now, since this system is nonlinear due to this nonlinear term, you can see that this term is nonlinear with respect to the control, and for that reason, you have to verify that this term is locally Lipschitz. So, how do we verify this? Term-wise, I am going to show you that it satisfies the Lipschitz condition.

So overall, it becomes Lipschitz. Okay, so now here in the first term I have taken, and then you can see that if

$$u \in \mathcal{U}$$

and $y \in \mathcal{Y}$, then I will be able to calculate some kind of Lipschitz constant such that this relation is satisfied. Similarly, for the term u^2, I now proceed by assuming that if I select the Lipschitz constant as 2μ, then this term also becomes locally Lipschitz. However, the Lipschitz constant basically depends on the region where $u_1$and u_2 are going to lie, and due to that region, the function is locally Lipschitz.

What is the meaning of globally Lipschitz? If the constant Lin the Lipschitz definition is independent of $u$and y, then we can say that the function is globally Lipschitz. So, in this way, this condition is satisfied.

Now, I am able to proceed towards the observer design. Since this dynamics is exactly known, I am assuming that the nonlinear dynamics is exactly known, because $u$and $y$are known to us. That is the primitive assumption whenever we are actually designing the observer, because we are assuming that $u$and $y$are available, and then I have to estimate $x(t)$. Due to that reason, this term is known to us. Now, since this term is known to us, I have to place the eigenvalues of $A - HC$ at some negative values.

So, I have placed the eigenvalues at $-3$and -4like this, and then you can see that this gives our error dynamics. So, I have defined the error dynamics as
$e = x - \hat{x}$.

So, in the literature, the error dynamics are defined either using $\tilde{x}$or using e. So, whichever notation you are comfortable with, you can proceed further. Now, I am going to take a more general class of systems.

You can see here that I am still assuming that the dynamics, that is, the nonlinear dynamics of the system, are well defined and well known to me. There is no uncertainty, but here I have a term that is called ϕ. So, ϕcontains those kinds of terms which actually appear in terms of $x$and u.

So, this term, the ψterm, I can assume is known. Why is that known? I have already discussed that since $u$and $y$are known during the observer design, this term is known to us. But the term $\phi$is unknown to us.

And now, whenever you proceed towards the observer design, you can see that your assumptions become more strict. What kind of assumption is that? So, up to here, the assumptions are exactly the same as the previous ones, but here $\phi(x,u)$should be globally Lipschitz in $x$and uniformly in u, because the user can select any kind of u, and for that reason, the condition must hold for all selections of u. You have to show that this function $\phi$is globally Lipschitz. It means that whatever Lipschitz constant Lcomes into the picture, it is not going to depend on the selection of $x$or u. Obviously, it may depend on x, but not on u, because the condition is uniform in u.

So, for all u, and due to that reason, $u$is not going to come into the picture in the Lipschitz constant, and this constant L should be independent of both x and u. Basically, it should not depend on either x or u. So, these are the kinds of conditions we have to maintain in order to show global Lipschitzness. What is the meaning of local Lipschitzness? In that case, the constant L can depend on the selection of $x$and u.

At that time, I am not able to use the word uniform. But here, in order to prove stability, I have to check that this term, whatever term contains the information of the state as well as the control, should satisfy the global Lipschitz condition. So, under that assumption, you can now see that the observer design is exactly the same.

What am I going to do? Since I have a mathematical model of the form
$\dot{x} = Ax + \phi(x,u) + \psi(y,u)$,
these two terms are completely known in the sense that $u$and $y$are available. Due to that reason, as it is, I have kept $\psi(y,u)$here. But here, the term ϕcontains x, and since $x$is not known to us, I replace it by $\hat{x}$. Similarly, I have $A\hat{x}$, and this is the correction term, exactly the same philosophy as in the Luenberger observer.

So, you can also treat this as an extension of the Luenberger observer. Now, here I am going to define the error dynamics using $\tilde{x}$. Now, you can see that $\dot{\tilde{x}}$is actually nothing but this term. Now, an extra term comes into the picture because I have some extra dynamics. In this dynamics, the term $x$is replaced by $\hat{x}$, which means that I cannot assume that it is completely known to us. But I am assuming that the mathematical model is known; I just do not know all the state information, and for that reason, some mismatches initially come into the picture..

So, this is nothing but some kind of dynamics where the linear term contains some kind of uncertainty. And this term, which I have already assumed to be uniform with respect to u, can be bounded by a constant multiplied by $x - \hat{x}$, and $x - \hat{x}$is nothing but $\tilde{x}$. And due to that reason, you can assume that this perturbation is some kind of vanishing perturbation. So, now if this perturbation is vanishing, it means that when the error is equal to 0, at that time there is no perturbation. So, you can expect that $\tilde{x}$tends towards 0as t tends towards infinity.

So, for that, I have to design H. So, what is our methodology? You can see that this system, even though it has a linear structure, is nonlinear in nature due to the additional terms. Due to that reason, I am going to proceed using the Lyapunov approach. Now, in the Lyapunov construction, the error is nothing but $x - \hat{x}$. Here, $P$is a positive symmetric definite matrix. What is the meaning of a positive symmetric definite matrix? It means that

$$P = P^\top$$

and all eigenvalues of this matrix are strictly greater than 0.

So, you can easily check that

$$\tilde{x}^\top P \tilde{x} > 0$$

for all $\tilde{x} \neq 0$, and it is equal to 0only when $\tilde{x} = 0$.

So, in this way, we can justify that this is a valid selection of the Lyapunov function. Again, whenever we proceed using this Lyapunov function, and if I have a linear-like system, it is possible to show that a Lyapunov matrix inequality of this form will appear, where this matrix is Q. So, instead of Q, it is possible to show that one can select $-Q$.

So, instead of Q, you can select $-I$. This is one of the standard selections. If the matrix

$$A - HC$$

is Hurwitz, then there always exists a unique matrix P. So, once you fix the matrix $A - HC$, then Pbecomes unique because $Q$is already fixed. So, in this way, this Lyapunov equation has a well-known solution.

Now, what am I going to do? I am going to take the derivative of the Lyapunov function, and I am going to substitute the error dynamics into the Lyapunov function. After that, I am going to apply the criteria of global boundedness and global Lipschitzness, uniformly in u, with the help of these terms.

And at that time, I will have dynamics like this. I hope that you are able to understand, because this is the identity matrix. So, $x^\top x$ is nothing but $\| x \|^2$.

And here, L is the Lipschitz constant, and after that, I have applied the norm inequality. Now, I have to make sure that this term is less than 0. So, how do you make it? You can see that the gain L should satisfy $L < \frac{1}{2\|P\|}$.

So, this kind of condition should be satisfied. If this condition is satisfied, how can one basically obtain it? The gain Hcan be obtained as a solution of the Lyapunov equation, and then one can prove that $\tilde{x}(t) \to 0$ as $t \to \infty$.

In this way, I can design an observer for the nonlinear system, and this kind of stability is global exponential stability. So, with exponential speed, I am able to drive the error to 0. Now, I have taken another system like this, and here you can see that the x, u term contains some kind of x^2.

So, there is no u here. So, obviously this system is uniformly Lipschitz with respect to u. You are free to apply any kind of u. Now, what I have done is designed $H = 3$, the same as in the previous example, and after that, I checked the pole or eigenvalue of this particular matrix, which is strictly negative. Now, I have to check the term $\phi(x, u)$, that is, whether it is globally Lipschitz. What is the meaning of globally Lipschitz? Here, I take $L = 1$, because for $\sin x$, one can show that $\| \sin x \| \leq | x |$.

So, I have used this particular inequality to obtain the Lipschitz property, and then, with the help of this Lyapunov function, I can design the gain just by solving the Lyapunov equation. Simply, what you can do is select Hin such a way that the eigenvalues of the matrix $A - HC$are negative. So, in both ways, you are able to proceed, but you have to verify this kind of condition. Even if you can place the eigenvalues, at that time you still have to verify this assumption. Then you can prove that $\tilde{x}(t) \to 0$ as $t \to \infty$, and consequently $x(t) \to \hat{x}(t)$.

So, in this way, you can be able to estimate the state. Now, it is time to conclude this lecture. So, what have we observed in this lecture? We have extended the framework of the Levenberg observer for the non-linear system, and particularly for whatever non-linear system I have taken, I am assuming that it is observable, and I have always considered some kind of specific form of the non-linear system. After that, I extended it for general form. Everywhere, what is our assumption? That mathematical model is known. But we are in this world where we have several different kinds of uncertainty.

So, it is not a very realistic assumption. that everything is known. All parameters, everything is known. So, always some kind of perturbation, some kind of uncertainty comes into the picture, even if I have a linear model or a non-linear model. And due to that reason, I need one more lecture to address that issue. So, for that, I am going to explore the high-gain observer in the next class, which is very close to the sliding mode observer. So, in this class, at least we are able to understand that the Leuenberger framework can be extended for non-linear systems provided some specific kind of assumption is actually verified.

What is that assumption? In terms of ψsystems, this term should be locally Lipschitz, and in terms of $\phi(x, u)$, this nonlinear term should be globally Lipschitz. Then, using Lyapunov analysis, we can tune the gain such that whatever mismatch exists between $x$and $\hat{x}$is minimized. And as $t \to \infty$, I can ensure that $x$converges towards $\hat{x}$exponentially. After that, I have taken at least two examples, and these examples are numerical examples. So, using those numerical examples, we have shown how one can design an observer in a systematic way.

I have also discussed gain selection, and I have done stability verification using the Lyapunov method. Now it is your turn. You can take any non-linear system and design the same kind of observer for the non-linear system. So, with this remark, I am going to end this lecture. Thank you very much.