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

Lecture – 40
Predictability A stochastic view and Summary

In the last module we talked about predictability in the context of deterministic model, in general deterministic forecast do not have any spread, deterministic forecast is always a point prediction at what time at I what day time of a particular year the lunar sonar eclipse will happen. So, it is a point estimate we are going to talk about the occurrence of a phenomenon called solar eclipse or lunar eclipse by specifying the year the month the day and the time.

So, there is no spread there is no variance, so in the context of solar eclipse it is perfectly predictable; we use deterministic models to be able to make weather predictions. So, weather predictions in general have always been point estimate until recently, but we know the point estimate by itself does not mean much we need to be able to tell the confidence with which the point estimate would be a reasonable, predictor this relation to the quality of prediction is often given by the variance associated with the value being predicted, in the case of deterministic models there is no notion of variance within the deterministic concept.

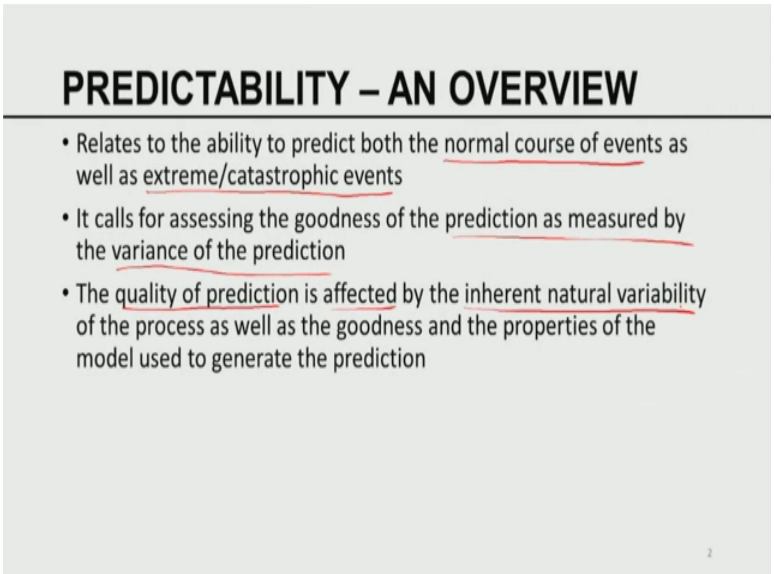
So, we worried about the era, I would have obtained one level of prediction had I used, one state I would have obtained another level of prediction had I used a state, which is closed by I would like to be able to look at the differences between this prediction resulting from a small minute of differences in the initial conditions, if the errors in the initial conditions grow with time then the model exhibits extreme sensitivity, the predictability does not predict I cannot make reliable prediction for long durations of time, in which case I have to determine the length of the period of time over which I have some confidence in the quality of prediction that led to the national predictability limit.

So, predictability limit is a measure of the time period within which I can believe I can I have a good idea of the goodness of the prediction, but the interesting in prediction is moving towards stochastic analysis, what is that tomorrow I am likely to have 10 centimeters of rain with the confidence of plus or minus 20 percent. So, that it is going to rain is for sure. So, what is that confidence on and the measure the predictable the level 10 centimeter rain; yes centimeter 10 centimeter of rain in 20 hours is an extreme event, we are all observing several extreme events over the whole world over a longer of time.

So, 1 of the problems here is that, what are the probabilities of extreme events? If you make a prediction on extreme event what is the confidence interval, what is the variance of the estimate? If I say I am going to have a prediction of 5 millimeter rain, what is going with the confidence of that? If I say tomorrows temperature is going to be 105 degrees what is going to be the confidence in here. So, as we move from deterministic prediction to some aspect of probabilistic prediction where, I am interested not only in predicting the level, but also the variability the uncertainty in the prediction that is the kind of the stochastic prediction.

So, motivated by this need for both the level and they spread of uncertainty which is characterized by variance in this module, we are going to look at some of the basic issues relating to predictability or stochastic view.

(Refer Slide Time: 04:48)



So, predictability relates to the ability to predict both the normal course of events as well as extreme or catastrophic events, is tomorrows weather is going to be like today's weather is there likely to be a chances for extreme events such as 10 centimeter rain in

24 hours these are extreme events. So, we would like to be able to predict not only the normal course of events but also extreme events.

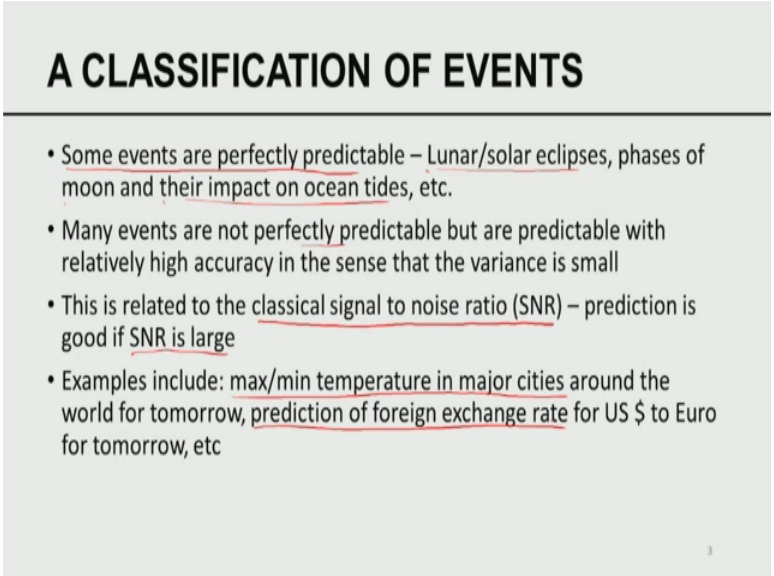
So, predictability problem has just dual goal normal and extreme, so it calls for assessing the goodness of prediction as measured by the variance of the prediction, what is the variance of the prediction of normal events? What is the variance in prediction of extreme events? Sometime we may be able to predict the normal events with better accuracy then extreme events, it is the extreme event that generally catches away by surprise, but these predictions are going to be created by models.

So, how do we calculate the level as well as the variance of these predictions, a proper appropriate framework would be to make a probabilistic analysis this probabilistic analysis is called the probabilistic view at the stochastic view of predictability. The quality of prediction often times is affected by both the inherent natural variability as well as the goodness in properties of the model used to generate the prediction.

We all know climate has a natural variability, 1 of the interesting climate sign is to be able to make how the temperature of the earth has been rising and what is the maximum raise that 1 would face in next 50 years, what is the effect of increasing carbon sequestration in the atmosphere on the raise of global temperature these are very interesting and practical problems; if you look at the predictions for the rise in temperature in the next 50 years 100 years vary anywhere from 2 degrees to 6 degrees, this there is a wide range of variation in this prediction that relates to the properties of the model that are being used.

So, different models have different characteristics leading to different types of prediction. So, ultimately the quality of prediction rust at the feet in the model, why it is a model that is used to extrapolate extrapolation is prediction. So, with this in mind we have already talked about the role of model in deterministic predictability. So, one measure of the deterministic predictability using predictabilities in deterministic model is the computation of lyapunov indices; if one of the lyapunov indices is positive then there is going to be a problem in trying to make long range predictions.

(Refer Slide Time: 08:21)



In the case of stochastic analysis, we need to be able to analyze what are the stochastic factors that affect the prediction that is what the ultimate role of this module. But before that we would like to classify some events some meaning full way some events are perfectly predictable we have alluded to this several times, but I think this contract is worth emphasize as some events are work perfectly predictable lunar solar eclipses, the phases of moon and their impact on ocean tides.

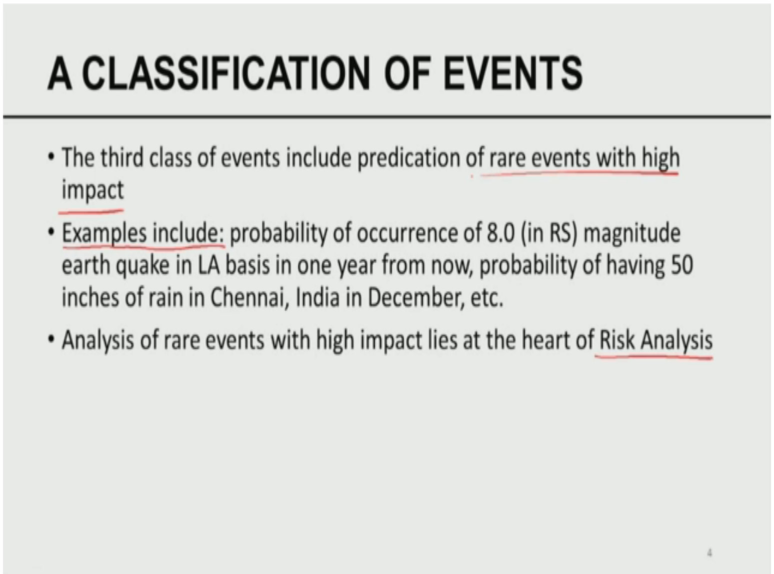
So, normally we make predictions about ocean tide at days during the week month, that depends on the phases of moon and those predictions are pretty accurate; many events some events are perfectly predictable but many events are not perfectly predictable. What do you mean by not perfectly predictable because, when you go from deterministic to stochastic the variance matters if the variance in the prediction is large then the quality of prediction is low, if the variance of prediction is low the quality of prediction is high.

So, in general within the stochastic framework this is related to the classical signal to noise ratio. So, a prediction is good if the signal to noise ratio is large, why that is the numerator signal denominator is noise, if the noise is low the signal to noise ratio is high. So, what is that affect that if that is reducing the quality of prediction is the noises, the noise can be thought of as errors in prediction those errors arise because of the property of the models that are used to create the prediction.

So, you can think of this we have alluded to this in the last lecture itself with respect to when we discussing when we are discussing the they Rayleigh's coefficients. So, Rayleigh's coefficients in some sense can be thought of as an indirect measure of the signal to noise ratio. So, what are the examples of normal events and other events, the maximum minimum temperature in major cities around the world tomorrow; the prediction of foreign exchange rate from US dollar to Euro for example, because for the contact of the economy I need to know the foreign exchange rate everyday foreign exchange rate is a random process.

So, we need to be able to and that depends on several factors affecting different currencies in different parts of the world and different economic conditions. So, there are very good mathematical empirical mathematical model they have been developed to be able to predict the foreign exchange, the foreign exchange a slight increase or decrease in rate means a enormous economical implications when you buy and sell. So, these are some of the normal events maximum temperature minimum temperature prediction of foreign exchange so and so far.

(Refer Slide Time: 11:41)



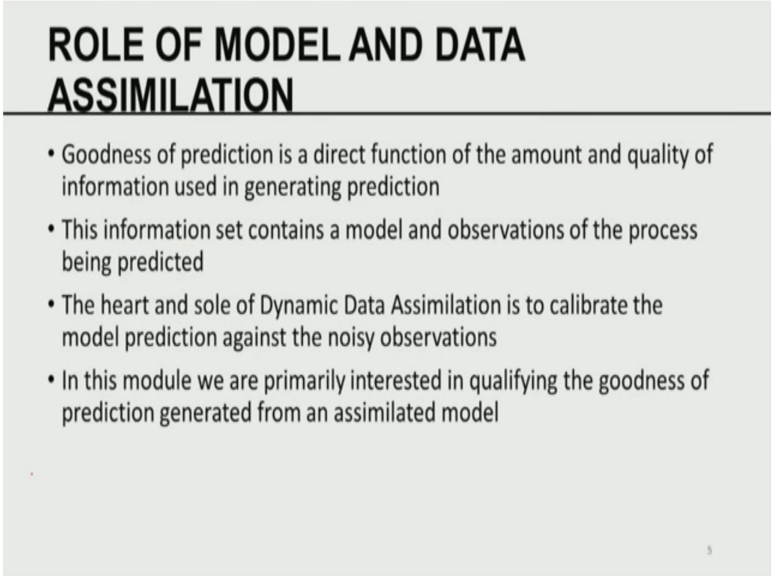
The next class of events is the prediction of rare events, rare events with the high impact for example, we recently witnessed a high impact rain event that effected certain parts of the state Tamilnadu it is a rare in the sense, it is considered to be a once in a 100 year event the impact is very high because, of continuous rain for several days that led to heavy flooding that cost immeasurable measure to several segments of the community. So, what are examples of prediction of rare events what is the probabilities that 8.0 in the Richter scale magnitude, earthquake will occur in the los angels basin in 1 year from now.

What is the probability of having 50 inches of rain Chennai in the month of December that is a rare event, analysis of rare event with high impact lies at the heart of what is called risk analysis for example, insurance industry are always interested in risk analysis. For example, we tried to live close to the water's edge, so there is a lot of population density all around the world along the shorelines, but shorelines are always subjected to extreme weather because of hurricane like events.

So, when you hurricane when a disaster hurricane comes and hits there is not only loss life loss of property. So, insurance industry when it is subjected to heavy loss of both life and property, they have to pay back lots of compensation and so they would like to be able to estimate the probability with which high impact, low probability when it will occur and based on which they will assess the risk to be able to decide the premium for insurance in order to be able to cover a mansion on the shorelines all over the world.

So, predictability relates to our prediction in general relates to normally when as well as rare events, in the case of high impact rare events it is more important to be able to make a good estimates about the prediction because, enormous implications in both loss of human and animal life plus loss of property. So, these are all the motivations for trying to improve the quality of prediction and also understand the level as well as the variability.

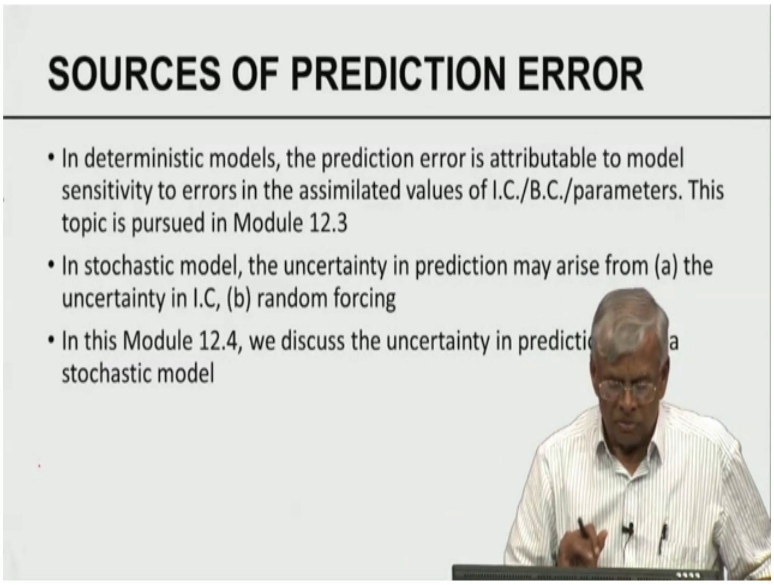
(Refer Slide Time: 14:33)



It is the general information that goodness of prediction is a direct function of the amount and quality of information used in generating prediction, this information set generally contains the model and the observations of the process being predicted, the heart and soul of dynamic data assimilation is to calibrate the model prediction against the noisy observations that what the primary emphasis is in this course has been, in this module we are primarily interested in qualifying the goodness of prediction generated by an assimilated model; an assimilated model means what I have estimated the unknowns the estimates have errors, the errors in the estimate induces errors in the forecast.

So, I would like to be able to knowing their statistical properties of the errors that leads to the prediction, I would like to be able to derive the statistical properties of prediction all within the stochastic context.

(Refer Slide Time: 15:39)

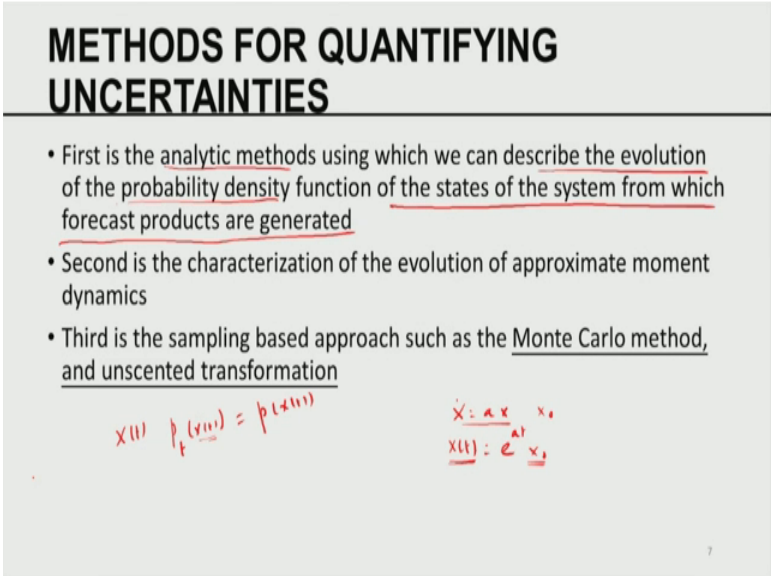


So, sources of prediction error there could be a some slight reputation that is I think is worth repeating some of the key factors, in deterministic model prediction error is a attributable to model sensitivity to errors; it is related to errors in the initial condition boundary condition parameters that are being estimated, it is the topic that we have pursued in the previous module.

In the stochastic model however the uncertainty in the prediction may arise from uncertainty in the initial conditions, uncertainty in the random forcing are uncertainty in the parameters; I am assuming we are assuming in here the parameters are very well established. So, I am simply trying to relate the uncertainty in the forecast to uncertainty in the initial condition and uncertainty in the random forcing.

So, in this module we have discussed the result to be uncertainty in the prediction induced by the uncertainty in the initial condition and are uncertainty in the forcing. So, in a deterministic model if the initial condition is random the solution exhibits random behavior, in a deterministic model if the forcing is random the solution exhibits random behavior, in a deterministic model both the initial condition and the forcing are random then the combination of the 2 results also in stochastic nature of the connection.

(Refer Slide Time: 17:18)



So, within this stochastic model setup we need to be worry about the uncertainties in the prediction induced by these 2 sources of randomness, the third source of randomness which we are not considering for the sake of simplicity is randomness in the solution induced by the parameters, they are often more difficult because their dependency of the solution and parameters are always non-linear more often, they are not even if a model is linear the solution depends on the parameter nonlinearly.

What is an example I will give a quick example, suppose x dot is equal to ax is the parameter that is the initial condition, we already know the solution x of t is e to the power of a t of x naught. So, inhere the if x naught is stochastic x t is stochastic, but x t depends on x naught linearly; however, if a the parameter is stochastic x t induces that induces stochasticity in x t, but x t depends on a of t non-linearly. So, even though the equation is linear sometimes the solution may depend on the parameter non-linearly. So, herein lies the whole spectrum of possibilities, the randomness can arise from initial condition can arise from parameters can arise from forcing here I am controlling unforced model.

So, what are the class of techniques that are available for analyzing solution to stochastic system and trying to compute or quantify the uncertainty into the solution that is the goal of predictability within the context of stochastic system? First is a class of analytical methods in this class of analytical methods, our aim is to be able describe the evolution of probability density function of this state's of the system for in which forecast products can be generated.

For example, if x t is a solution of the system at time t if p t of x t is the probability density of the solution at time t we know p t evolves in time. So, there are 2 t now x depends on t p is the density function of x of t the density function of x of t may also depend on t because p t evolves in time, if p t is equal to is equal to p of x of t what is it mean the probability density does not change in time that is what is called invariant density are stationary density; if x t varies accordingly to the dynamical system it is often the case that p also varies. So, we are in general concerned with we are in general concerned with p t of x of t so probability density.

So, if there is a means by which I can describe the dynamics of p t of x of t based on the dynamics of x of t, what is that we know we know the complete distribution of the state of a system at every time; what is the maximum information one can give about random variable is probability distribution, there is no more information than the probability distribution that one can endowed a random variable with.

So, once the probability distribution is known we know everything about it, once given the probability distribution I can complete the mean variance I can compute any number of status statistics you want to be able to generate out of it. So, knowing the p the probability density function say what are the forecast products mean is the level variance is the uncertainty. So, I can create the forecast products as mean and variance from p f t, if p f t can be analytically derived, but it turns out that the ability to characterize p t analytically is possible only for a very small class of uninteresting simple textbook cases, so that is the difficulty.

What is the second one the second one is if I cannot do everything? So, knowing p t is everything if you cannot do everything at least I would like to do say meaningful pacts of everything. So, what is that that gives rise to approximation of moment, we have already talked about in the contacts of non-linear filters how do we generate moment approximation in the case of linear dynamics in the case of non-linear dynamics. So, that is the next best thing evolution of moment dynamics, the evolution of evolution of moments or moment dynamics.

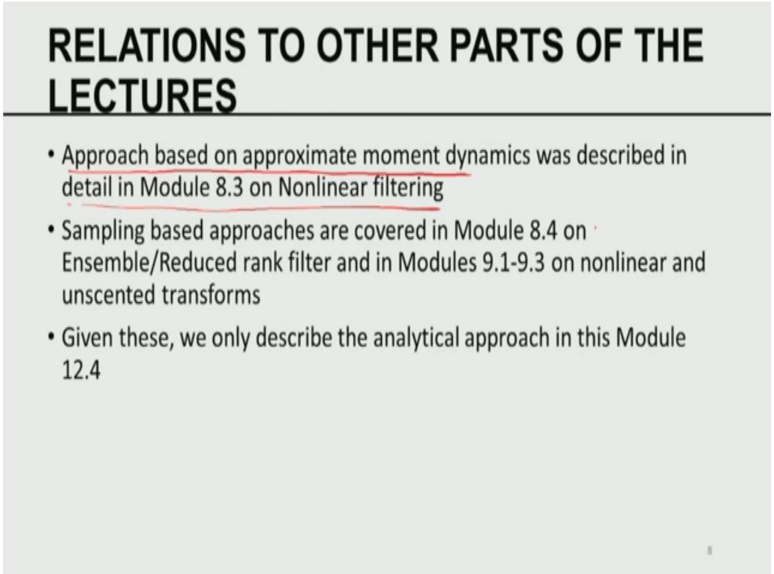
What are the third methods? Third method is called a Monte Carlo technique, what are the Monte Carlo techniques. So, if I am given a deterministic system if the deterministic system has initial condition random then, I can pick samples from the initial distribution pick different initial conditions from the given probability distribution run the model forward in time, I can create an ensemble of forecast and that is the basis for ensemble method that we have seen in the context of ensemble filters or reduced rank filters a class of reduced rank filters coming from ensemble methods.

So, once I can create an ensemble of forecast any at every time then, I can compute the ensemble mean I can compute ensemble variance or covariance, I can use the ensemble mean and the covariance at the forecast product. So, these are primarily the 3 ways of attacking stochastic predictability issues.

Third method Monte Carlo method is useful now in all cases, but it requires lot of computational power because I need to be able to run the model in parallel for emptying the different types of initial conditions, how many such initial conditions you should be able to run in order to able to at meaningful estimates of the variance and covariance, that comes from large sample statistical analysis.

So, statistic provides you in order to be able to estimate the mean with certain degree of accuracy how many samples you must have, in order to be able to estimate the variance with certain degree of accuracy; how many samples you must have? So, we can fall back on the large sample statistical theory in trying to estimate how many ensemble members do I mean to be able to make descent estimates of the mean covariance analysis based on which we can generate forecast products for public consumption.

(Refer Slide Time: 24:08)

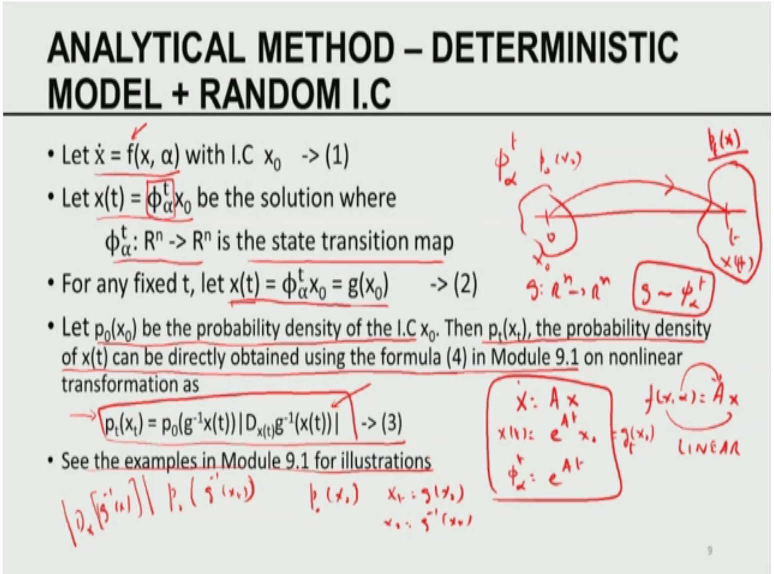


So, relations to other parts of the system, I would like to now talk about this the approach based on approximate moment I have already allowed you to, but we would like to formalize it an approach based on approximation moment dynamics was described in the module 8.3 on non-linear filtering. So, even though we are not covering it here I will many parts of this idea have already been covered.

So, I would like you to be able to tap those results to the benefit of making probabilistic prediction, sampling based approach which is the Monte Carlo approach have the it has was can be related to ensemble reduced rank filters and that was covered in one of the earlier slides, there are other class of methods called non-linear unscented transformation we are not going to be talking about, we will not have time to cover these and there are enormous literature are on this area.

So, unscented transformation is a class of ensemble is a specific class of ensemble methods using which you generate not only the stochastic prediction, but also it will enable you to be able to compute the moment statistics rather easily. So, given these that we have covered aspects of Monte Carlo methods aspects of moment dynamics earlier, aspects of Monte Carlo coming from ensemble aspects of moment dynamics in the context of ensemble filters, I am simply going to restrict our discussion to the class of analytical approach; admittedly this approach is very limited because it is solution processes except for a simple cases I cannot solve this even though it is called analytical approach, the equations for the evolution of the probability densities are clearly known that the solution closed form solution is hard to combine, but using those solutions 1 could use numerical methods to be able to develop good forecast and quality of forecast using computing variances.

(Refer Slide Time: 26:27)



So, we narrow down all our discussions to one class of method namely an analytical method. So, what is the condition under which one class of analytical methods can be applied the model is deterministic, the randomness in the solution comes essentially from randomness in initial condition, so that is the first case. So, let us take for simplicity I am now dealing with the continuous time there is a corresponding analog of this treatment using discrete times.

So, let is start with an ordinary distribution coefficient, let x dot is equal to f of x comma alpha with an initial condition x naught alpha is the parameter I will let is assumed alpha is well known alpha is error free; if this is the differential equation dynamical system clearly tells you the solution at time t can be described in principle by a map, the map is phi t of alpha. So, phi alpha t is a map r to r n this map is called the state transition map, this map tries to relate that initial condition t is equal to 0 to the solution at t, so it relates x naught to x of t. So, relating x of t directly to x naught is done by the use of a map this for a fixed x t, the solution x of t is simply a function ac ah of x naught. So, phi f t of alpha of x naught is called it can be can be relabeled as g of x naught.

So, if the emphasis in here what is the emphasis in here x t is a solution of the system under mild conditions and f such as condition it can be shown that the solution exists solution is unique, if the solution exists on unique there is a representation for the solution. So, the solution can be represented by a transformation of the initial condition, the transformation is just phi of alpha t and that is what is called the state transition map.

For example if x dot is equal to A of x in which case f in other words f of x of alpha is equal to A of x the elements of alpha may be in matrix A x is x this is linear, in this case the solution x of t is e to the power of A t times x naught, therefore in this case phi t of alpha is essentially e to the power of A of t. So, phi of t is not something unknown in the case of linear system I can quantify what that map is, but in general under mild conditions on the existence of the solution; exist means I should be able to relate x t to x naught the relation between x t and x naught is defined by the transformation or a mapping and that is what is called the state transition map.

Let p naught x naught be the probability density function of the initial state x naught pt xt be the probability density of x t, that can be directly applied using the formula 4 in module 9.1, unfortunately we are not covering 9.1. So, I am going to give you the formula here. So, you do not have to worry about 9.1, but they the formula that is referred to is given by the relation 3 in here what is that? We have already said that x of t is relation is a function of the initial condition, I can think of x of t to be g of x naught.

So, g is the mapping from r n to r n and g is a related to phi of alpha of t g is relative alpha of alpha of t for alpha of t. So, in the case of a linear once again x naught is equal to a x t is equal e to the power of a t x naught. So, I can relate this as to g of x naught in this case as g t of x naught because, e to the power of a e to the power of a t depends on t, so I can simply say if g t of x naught.

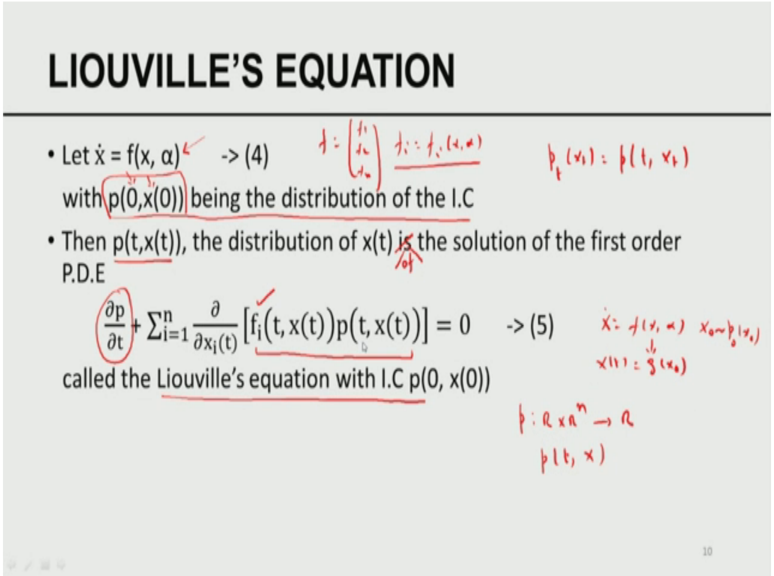
So, now let us consider of the relation it can be assumed that this initial condition is random with the probability p naught of x naught, if the state of the system at time t the probability density p of x of t, it is a simple exercise in probability theory it is simply transformation of random variables g is a function that leads to the transformation. So, I can express the density at time t in terms of the density at time, 0 times a multiplying factor that is the magnification factor.

Now, let us look at the 5, let us look at what is happening in here. So, p naught is the initial density that is evaluated as g inverse of x of t; why g inverse of x of t because, the distribution of x naught is known g of x g of x naught is the value of the solution at time t, so g maps the initial state to the state x t. So, under mild conditions on the existence of solution it can be shown. In fact, I am not proving this we have already proved this is already proved in several text books and probability theory it is called transformation of random variables or transformation of random vectors. So, the formula 3 essentially captures what we want.

So, if I know the initial distribution p naught of x naught, by doing an inverse transformation so if x of t is equal to g of x naught, so x naught is equal to g inverse of x of t. So, p inverse p of 0evaluated g inverse of x of t times, what is the multiplying factor it is the determinant of the Jacobean, so b x of g inverse of x. So, g is a function g inverse is a function Jacobean of that function is a matrix that determinant of the Jacobean of that. So, that is called the magnification factor if this result is very well known in probability literature. So, I can compute the distribution of a random variable of the solution at time t in terms of the initial distribution and the magnification factor.

There are number of examples 1 can used to illustrate this, these examples are very well known given in number of text books in basic course in probability theory; I would refer the reader to some of these good books for example, fellers book is a good example Papoulis books is a good example. So, there are number of interesting text book that relates to the transformation of probabilities, so 3 essentially gives is analytical solution.

(Refer Slide Time: 34:20)



So, to be able to use this equation 3 what is that I should know, I should know the solution, I should know in other words I should be able to I should be able to solve the system. So, look at this now I should be able to solve this system grow as the same problem as I had just morning I am sorry.

Student: Right sir.

See I defined all these things in the context of in the context of uncertain transformation I have what the lot of examples. So, I can I have what is I am doing wrong when it sees you it behaves better than that know.

Student: Sir.

One second go thank you.

Student: It is correct sir.

So, what is that we have done, we assumed the basic things like this I have an equation I have a dynamical equation x dot is equal to f of x of alpha, I have given initial condition is the distribution p naught of x naught; I know the solution x of t as g of x naught; that means, I know the solution explicitly the solution is the function of g. So, if you know the solution of the differential equation then the formula 3 in the previous page, essentially gives you the evolution gives you the distribution of solution x t at any t.

Now, I would like to be able to reflect on that linear system we can solve them perfectly because, x dot is equal to x is a linear system in the case of linear system the solution is always given by x of t is equal to e to the power of A of t of x naught. So, for all linear system this formula is extremely useful and simple, but there are very few non-linear system for which we can solve the system, there are very few non-linear system for which we know the g; please remember the expression in g is counting the expression in 3 conditioned on a knowing g, so if g is not known 3 is not useful.

So, in cases where so what is the summary, I know I have a differential equation I know my initial conditions are random, but I can solve the differential equation and get the solution in an closed form x of t is equal to g of x naught. So, once I know the g I can use the equation 2 in this module to be able to characterize the probability density function at time t using the formula 3, so that is the basic idea.

The question here is that what is if I am not able to solve. So, that is the next question and that is what being answered, in this slide we are going to assume I have a differential equation I have an initial condition that is random, here I am going to change my notation a little bit instead of saying p t of x of t, I am now going is going to say p t of x of t really does not matter; we repeat this time subscript as A subscript or as a form of along with the functional dependency.

So, p 0 x 0 what is it mean this 0 tells you initial mean this tells you the initial state. So, if I know the probability distribution of the initial state at time 0, I would like to be able to compute the probability distribution of the solution at time t. So, p t of x of t is the distribution of a x of t, distribution of the solution map is the of the first order p t that is given above in 4.

I am now going to drive this it can be shown this p t of x of t is given by the solution of a partial differential equation, where the partial of t with respect to t plus the summation from I is equal to 1 to n of the partial of the product f I versus p t now p is a scalar function please remember, p is defined on I am sorry p is defined I want to make sure the we understand this p as a function p is the probability density function, it is defined on r cross r n to r. So, p has a first component time second component x is a matrix and the value p of x is always a scalar, the probability is a number in between 0 and 1 we all know that and the integral of the probability must be equal to 1; f I is the I th component of f now go back now what is the f? f is equal to f 1 f 2 f n each of the f I are functions of x and alpha. So, each of the each component f I is a function of x and alpha.

So, I am going to consider the product of f I with p, I am going to sum with respect to I, I am going to consider to the partial with respect to x I of the product, what does this say the time derivative that the derivative of p with respect to time plus this term must be equal to 0 this equation is called liouvilles equation; this equation essentially tells you the conservation of the probability math's. So, the probability density function may vary in time, but what are the conditions it must always be non it must be always non negative and the total amount of probability must be preserved equal to 1, this is a simple statement about the conservation of the probability math's at all time that leads to this partial differential equation.

Now, please understand fi is are known p I p t x t are not known what are the claim the solution pt xt is obtained by solving then the partial differential equation in 5, this equation is called liouvilles equation; I am not going to indulge in the derivation of this then derivation itself is a very good exercise their lots text books in stochastic dynamics, that talk about liouvilles equations and solution process one of the standard ways of solving the liouvilles equation is by method of characteristics.

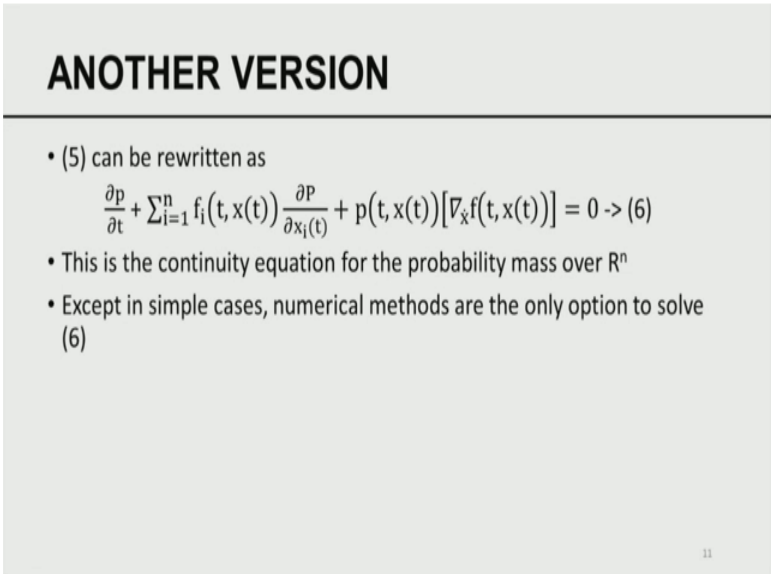
Those of you have studied partial differential equation theory should be very familiar with method of characteristics. So, you can solve this equations very readily there are lots of good text book, that deals this the solution process of first order equations by method of characteristics. So, fi is known pi is the unknown, pi is a obtained by solution of this partial differential equation.

Solving this numerically would be an would be a very difficult task because, you have to discredited you have to compute this you have to ensure the probability is never less than 0, the probability value is never ah total sum must be equal to 1. So, to the preservation of their probability mass at all the time leads to lots of computational challenges, if one tries to solve this equation in numerical forms, this equation again can be solved in analytical form only for certain special classes f is again we are against the hard rock; we know how to do,, but we may not be able to accomplished because, the method of solving liouvilles equation analytically is very difficult, numerically is also very challenging, so these are some of the tools that are available. So, a liouvilles equation is one way by which we can characterize the evolution of probability density function of the solution of a differential equation whose initial condition is random.

Once suppose you are able to compute p of t a we know everything about x of t then we can compute the mean. we can compute the variance. we can compute all kinds of statistics if you want and that could be used at the forecast product. So, that is how uncertainty in the forecast within the context of stochastic dynamic system are quantified.

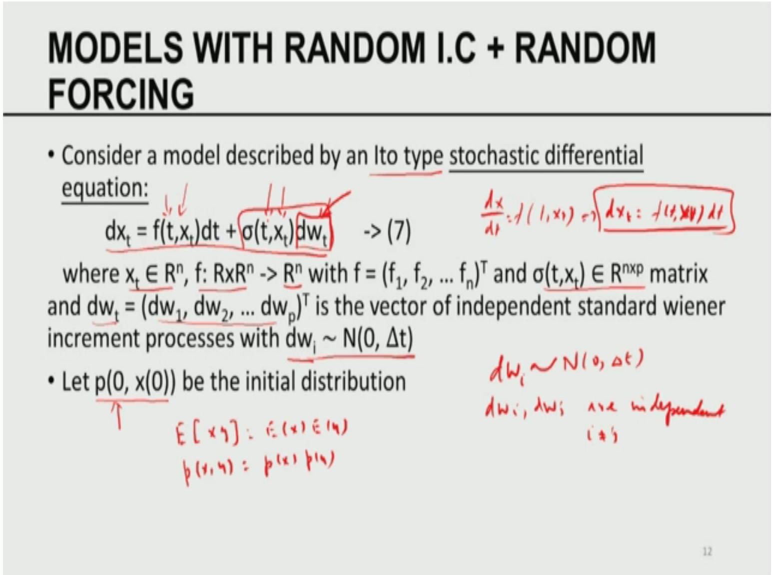
So, so far we have talked about uncertainty arising only from initial conditions. Again I have not shown you how to derive the formula 3, I have not shown you how to derive this equation 5, I believed these things could be a motivating factor for a for the reader to be able to engage himself or herself in different directions, I wish you will perceive different lines of enquiry depending on your interest. That is the way to rewrite this. So, I can do the differentiation on the second term and I can readily write this equation as equation 6.

(Refer Slide Time: 44:00)



This equation is called the continuity equation for the probability mass r of n, I have already I have already discussed that. Again I want to reemphasize acceptance simple cases numerical methods or the only option to solve 6. Even here designing numerical methods to solve these kinds of partial differential equation is extremely challenging the caste of the constraint of this you can solve numerically, but not every numerical solution will be able to satisfy the constraint of probability mass being preserved. So, your numerical methods have to embed this condition as a part of the design, and that is largely the challenge. Again there are books written on how to solve partial differential equations of this kind, there is a normal literature I would like you to I would encourage you to enquire into this line of a argument.

(Refer Slide Time: 45:02)



Now, I am going to go to the next level, where I am assuming there is a randomness in the initial condition plus randomness in forcing. There are 2 sources of randomness in this case a perfect way to be able to describe this stochastic differential equation is under the amble of what are called stochastic calculus defined by Ito. So, they are called Ito type stochastic differential equation. A typical Ito type of stochastic differential equation can be written in this form if sigma is 0, the second terms is 0 in which case it becomes d x by d t is equal to f of t of x of t, let me f of t of x of t. So, f of t of x of t now can be written their differential form as f of t of x of t of x of t times d t. So, this is the deterministic part. So, this is a random part, sigma t x t is a state dependent function, this is again a state dependent function this state dependent function can also explicitly depend on time, this function also can depend on state and time. So, sigma x is a could be a complex non-linear function of both time and state. F is also could be a complex non-linear function of both time and state, d w t this is the important new comer, dwt is called wiener increment process it is the very important class of stochastic process is.

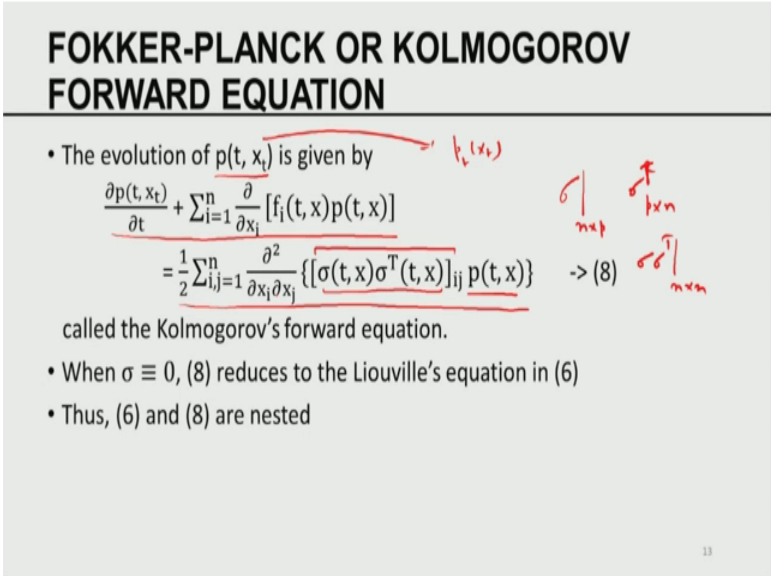
What is the property of the wiener increment process d w t in this case is a vector. So, before I consider d w t, let me talk about some of the dimensions of quantities involved x t is R n, f is the map depends on 2 quantities t and x n, t comes from r x and comes from r n the value is a map. So, it is in R n. So, f is equal to f 1 to f n, sigma t is in Rn by p matrix. So, w d t is a vector with p components w x t is a vector with b components. So, the ith component of this vector is called w d w i if I consider the ith component, I am going to talk about the properties of this, this is a random variable which is normally distributed whose variance is delta t what is delta t? Delta t is a short interval of time.

So, this kind of a stochastic process is what determines the first in term, this is called. So, w d w i and d w j are independence stochastic process is for i not equal to j. So, the vector d w t is a collection of p wiener incremental stochastic processes, which are linearly which are stochastically independent with the each other. Stochastic independence means what? What is the property of stochastic independence? If x and y are random variable such that they are stochastically independent, E of x y is equal to E of x times E of y.

If x and y are stochastically independent the joint density is the product of the marginal densities. So, you can talk about lots of properties of independence. So, independence of very strong property that using the stochastic independence of this p component vector, where each component is normally distributed whose variance is proportional to a small incremental time delta t, there is a special class of processes is called wiener processor. So, that is going to be the random forcing.

So, x of t is going to be the random because of the random forcing, and also x naught is random. The x naught is distributed according to this distribution. So, this is the initial distribution that is random, this is the forcing that is random, the forcing is magnified by the term sigma sigma. So, sigma is a kind of a matrix that tries to magnify the effect of forcing. P is the size of the random vector, the least value p can take is 1 in which case sigma t is the vector, and w d w t is the one standard increment. In the extreme case p could be n in which case there could be n linearly independent wiener increment process. So, with p as a variable between 1 on n, this equation 7 describes a finally, of stochastic differential equation stochastic ordinary differential equation. In general any stochastic model can be represented by this, every model of this type every equation of the style represents some form of stochastic models.

(Refer Slide Time: 50:26)



In this case I am not again going to prove this, this is been proved in several reasonable intermediatory level courses on stochastic processes. So, what is that our interest? Our interest is been trying to study p t of x of t, what is p t of x of t? The probability density of x of t at time t, the probability density this is what we had so far mentioned it as p sub t of x of t both are same. I would like to be able to derive an equation that governs the evolution of this probability density that equation is given by this system that is given in 8.

The left hand side of this is the same as in liouvilles equation, you can readily see that. The right hand side is the new term that comes in to be into being; the right hand side squarely depends on the matrix sigma. So, sigma is a matrix sigma transpose is the matrix please remember sigma is p by is n by p, sigma transpose I am sorry sigma transpose is p by n therefore, sigma sigma transpose is n by n. So, the this is. So, take the I jth element of this n by n matrix multiplied by p, and compute the second partial of this product with respect to x i and x j, sum overall i running from 1 to j, j running from 1 to n that is a right hand side. If sigma is 0 it reduces to a liouvilles equation, when sigma is 0 7 becomes deterministic differential equation.

In the deterministic differential equation stochastic stochasticity occurs only from the random initial condition. So, you can readily see the left hand side captures this stochasticity arising from the randomness initial condition; the right hand side captures the stochasticity coming from random forcing. The this equation is the very well known equation derived by Kolmogorov in the 1930s is called Kolmogorov forward equation. In some circles in physics, this kinds of equations were derived much earlier in the early 1900 starting from the days of Einstein who tried build the model for Brownian motion, Chapman, Kolmo, Fokker, Planck these are some of the names of names of associated with it. So, the some of the earliest name in the physics literature for this equation is called Fokker Planck equation.

It is Kolmogorov in the thirties he formalized everything by putting everything on a very strong mathematical basis. So, physicist had used lot of good intuition to come up with very good mathematics,, but Kolmogorov as he often does put everything on a beautiful mathematical pedestal where he developed the basic axiomatic approach to probability theory, he then developed the axiomatic approach to analysis stochastic processes in particular class of micro processes, and he tries to a characterize the solution for Markov processes. So, Chapman Kolmogorov equation is one form another is in forward chap Kolmogorov forward equation in is another equation that described that is related to Markov processes. in terms of Markov processes, the model 7 actually defines a continuous time Markov processes, we have already dealt with Markov process in the context of non-linear filtering in discrete time, and this is in continuous time.

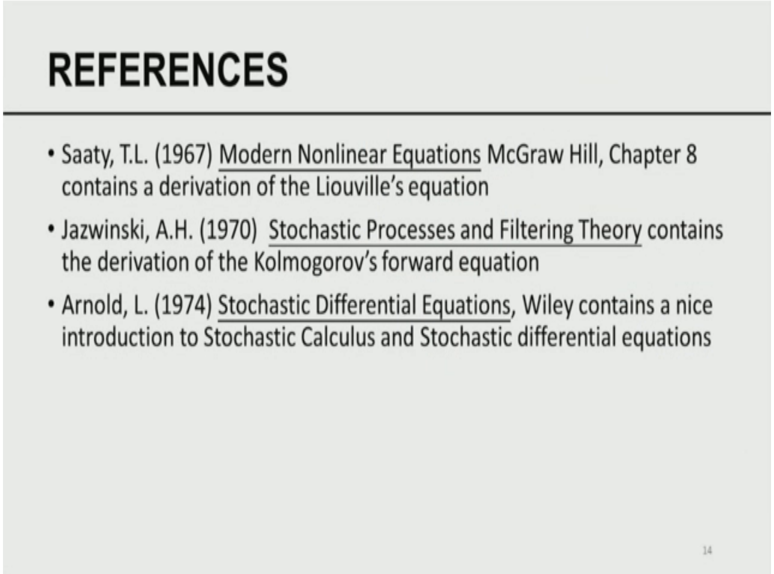
So, Kolmogorov forward equation essentially tells you how the probability density of a state function evolves in time, when the model is subjected to 2 types of randomness one coming from randomness initial conditions, another coming from randomness in the forcing function. The forcing function of the type induced by the wiener increment; you ask a question why you are you talking about particularly talking about randomness coming from wiener increments, that takes to for a for into stochastic modelling which we are not going to get into at this stage. So, a general description of a Markov model is given by 7 within this stochastic analysis framework. It is within this general framework Kolmogorov forward equation describes the evolution of the probability density function. So, you can readily see I have already already too, but it is worth of repeating it. When sigma is 0 8 reduces to Louvilles equation therefore, equation 6 and 8 are beautifully nested.

You can also see the addictively property that comes into be. If I had only initial condition what happens, if I had initial conditions and randomness in the forcing what happens? Unfortunately there is no general equation that talks about the evolution of probability density function, when there is stochasticity in the initial condition, stochasticity in the forcing, and stochasticity in the parameter. Analysis of stochasticity in the solution arising from stochasticity in the parameters, is extremely hard and that can be that is been done only for a very special cases. There was a call random differential equation with the random coefficients, that is lot of literature within the engineering community, but the methods are adhoc. There was no one grant the theory that combines initial condition uncertainty, forcing uncertainty, and parameter uncertainty.

So, uncertainty quantification is the primary goal of stochastic predictability analysis, if we make a prediction I need to be able to understand the uncertainty associated with the prediction. So, uncertainty analysis stochastic predictability they are all related disciplines the tools. So, we talked about some of the available tools, that 1 has at ones disposal to be able to get a handle on this stochastic predictability, again you can see solution of Kolmogorov forward equation is also an extremely difficult, because it is also a partial differential equation.

Again the mass conservation must be satisfied. Physicist have relied on Louvilles equation and Kolmogorov equation in the past, for simple systems they have solved it very cleverly in special cases. So, you can look into books that deals with stochastic processes in physics stochastic processes in physical sciences, that deals lots about stochastic model that has arise in the context of physical systems, especially of interest in basic physics and basic chemistry.

(Refer Slide Time: 57:44)



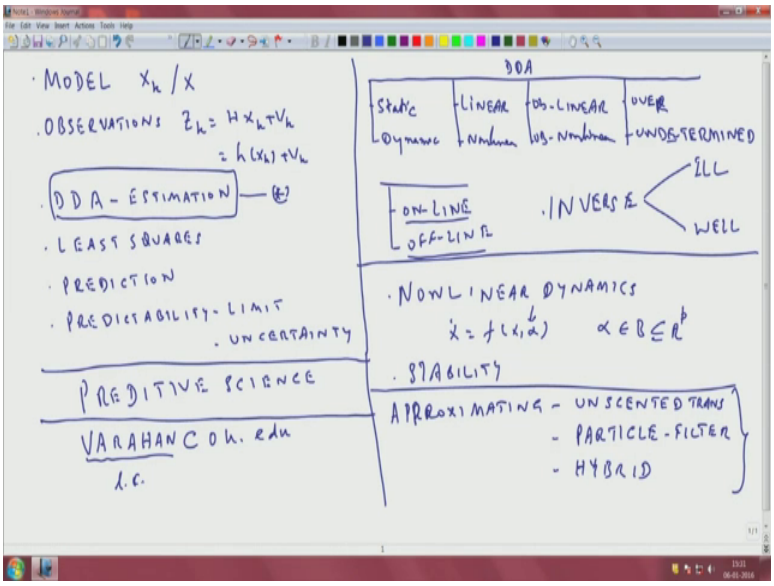
With this I believe we have come to the end of a discussion of the stochastic predictability. So, stochastic predictability is little bit more complex than the deterministic predictability, because quantifying the evolution of the continuous function p of t of x of t is an infinite dimensional problem. They are solutions of partial differential equation which are often extremely difficult to solve numerically, there are several attempts. The are reasonably good methods to solve these things, but they are all rather demanding.

Where we go for a good reading of this, Liouvilles equation is derived in chapter 8 Saaty's book 1967, modern non-linear equations is one of my favorites. Chapter 8 is a very succent chapter that provides a beautiful discussion on the derivation of Euler's equation. Jazwinkis book on stochastic processes and filtering theory contains a very readable derivation of Kolmogorovs forward equation as well as backward equation. Arnold; Lubrig Arnold nineteen seventy 4 stochastic differential equation published by wiley contains a very nice introduction to stochastic calculus and stochastic differential equation, that is also my favorite.

I use Arnolds book in my courses on stochastic differential equations, stochastic modelling and related topics. So, Saaty is very specific to the discussion of liouvilles equation, Jazwinki generally talks about the whole aspects of stochastic filtering. So, Kalman filter non-linear filter they are all discussed at great length in Jazwinkis book. He also describes in, I very good detail introduction to stochastic calculus from engineers perspective. So, Jazwinki would be a very good starting point for those us who are interested in pursuing stochastic modeling stochastic analysis. Of course, a much more rigorous mathematical treatment of this stochastic analysis is book by Arnold; with this we conclude our short introduction to stochastic predictability. We have not come to the end of our course relating to dynamic data assimilation, some going to provide a broad summary of what we have done where we are and where you can go.

Data assimilation problem: first you need to create models. So, model generation is a topic in itself. So, the modelers whose primary aim is to be able to develop models, for various physical phenomena of interest.

(Refer Slide Time: 60:37)



Then there are people who are measurement people, who develop methods for observations. They develop various sensors using which they can measure different quantities of interest to various disciplines, pressure, temperature, humidity, wind velocity and so on. So, observation people they are always interested, what is that we you can we are capable of observing directly. They can observe the energy radiated, they can observe a reflectivity from a cloud, certain observations are direct certain observations are indirect. So, model describes the state x k or x, x is a static model x k is a dynamic model, observations are z k, we have already seen observations are related to the model through a function h which is called a forward operator, observations are noisy this is a h of x k plus v k. Then once the models available, once the observations are available then we need to be able to calibrate the model the calibration process is called data assimilation dynamic data assimilation DDA. With the models categories simply call data assimilation, you know model is dynamic we call it dynamic data assimilation. So, data assimilation relates to estimation process, I am not going to repeat everything we have talked about the entire course was on estimation.

The unifying theme is the least squares method. Once we have gotten the estimation, then we want to be able to generate prediction. Once we are able to generate prediction I want to be able to talk about the quality of prediction, predictability limit. Predictability limit in the case of discrete time model uncertain, quantification of uncertainty in the case of stochastic model uncertainty in prediction; this is by enlarge our overall theme, be largely confront ourselves to the estimation prop to the estimation and data assimilation part of the whole story.

So, you can see how many different branches of science are involved modeling development of observational systems, then development of least squares, development of methods for estimation, ability to predict, ability to understand the quality of production. So, I would say these together all this together forms the basic components of predictive science.

So, predictive science is fundamental to many many aspects of our human life and within this predictive science. We concentrated on this process namely estimation process. So, we can find our attention to estimation process, but we classified our presentation across various topics DDA. We talked with model being static or dynamic, we talked with the model being linear or non-linear, we talked about observations being linear non-linear, we talked about the we talked about the observations linear non-linear and other things and we talked about the over determined case, we talked about the under determined case over determined or under determined.

In the over determined case what is the idea? I have more data than the number of parameters to be estimated. Is the dimension of the unknown to be estimated is n is the number of observations is m, m is greater than n that is called over determinant system I have lot more data. So in fact, let me (Refer Time: 65:34) satellite skip splitting data radar meteorology radar keep splitting data I have a lot more data than the number of variables I need to estimates. So, many of the problems in stochastic meteorology radar meteorology, generally lie in the framework of over determined system. In the under determined system, there is much less observations.

For example, I would like to be able to explore whether there is oil whether there is gold. So, how do they estimate the amount of gold in this mountain before I start escalating? They go and drill holes, whether it is for exploration of oil, whether it is to be exploration of platinum, where exploration of rare metals, they go and take samples from the earth by drilling holes from the samples. So, they get their drill tray number of holes for each hole they get the samples.

From the samples and use using the topography of the formation geological formation, somebody is going to estimate well, within these mountain containing this much of volume we estimate there will be 10 to the power of 5 kilos of gold with an estimate. Why do I need to know the an estimate of of the amount of the gold, because what is the current price of gold.

So, what is the net worth that gold that is buried in here may worth? So, if I know the amount of money that the that is hidden here is worth, then I can see how much money I can spend to dig it out and process it and make profit. How much money I can spend to be able to ah pump oil to be able to make money? So, all the oil companies in the world they want to be able to estimate where the oils are or where the oil reserves are, what is the amount of oil reserve estimate. And then based on the estimate and based on the cost of production, they try to decide to drill a hole here to drill a hole there online or of shore in shore and so on and so forth.

Now, to drill 1 hole for petroleum these days in the current technology, I was told it costs anywhere from 12 to 15 million dollars one whole; 1 whole of depth about 2 2 miles. So, how many 15000 dollars one can spend before they estimate? The amount of money is limited therefore, the number of observations that can be made in making predictions about natural resources hidden within the earth, that is an estimation problem there I have to make an estimation based on a smaller set of samples, because of the cost because of the simple shear cost of collecting data. In the case of satellite, a satellite may cost it 200 million dollars, but once you put the satellite once, it is going to be live for 20 years.

So, you take the money over time, you take the branches over time. So, depending on what is being measured, what is the cost of making measurements? A problem can be divided into over determined and under determined problem, also we have talked about online process off line. So, we have we have done various combinations, we are talked about static linear with observations are linear under determined over determined. We are talked about online off line, we are talked about dynamic linear non-linear observation linear non-linear.

We always generally talked about over determined system both offline and online. What are the specific class of methods, we have talked about the importance of deterministic least square formulation a stochastic static least formulation, dynamic least square formulations stochastic dynamically least square formulation, we are talked about on line methods, which are sequential methods it is a Kalman filtering, we are talked about off line methods such as 4 d var.

So, we are covered a 10 of topics, in addition to covering these kind of topics we have also emphasize the importance of mathematical tools from finite dimensional vector space matrix theory, optimization theory, theory of multi valued calculus, matrix algorithms as well as optimization algorithms. So, in about half the course we have emphasized the importance of thorough understanding of the fundamental mathematical principles that lie at the core of data assimilation why? Without these fundamental mathematical principles you cannot proceed to be able to do data assimilation.

So, while model building relates domains specific knowledge, where the models are going to be ah developed. Well observations also under to develop measurements of observation you will understand good physics, but relates to does not have sensors. In our view data assimilation is kind of an engineering discipline, we are interested always in trying to find optimal estimates of unknown. So, it leads to solving inverse problems. So, the kol key is inverse problem, and while trying to solve the inverse problem we have talked about ill condition problem as well as well condition problem. Generally inverse problems are hard to solve, a well condition inverse problem is easily solvable than ill conditioned inverse problems we have given instances thereof.

So, with this I think you get a good understanding of what this subject is all about. The sum is always greater than the parts, there are lots of parts to this story,, but the sum total provides you a broad overview of the area that has come to be called dynamic data assimilation now. So, the whole course provides introduction to data assimilation across these dimensions, static versus dynamic, linear versus non-linear model linear observation or non-linear observation over determined under determined off line versus online and all the mathematical prerequisites thereof.

Now, what is the next? Next is more detailed study of the predictive predictability studies. We have scratch the surface in 2 lectures. So, predictability study of deterministic system involves a deeper results from non-linear dynamics. So, non-linear dynamics I am sorry a non-linear dynamics, I misspelled it a good introduction to non-linear dynamics will help us understand, the variation of the solution with respect to changes and parameters. For example, x dot is equal to f of x comma alpha we assume the alpha is fixed. So, it behooves to ask a question what is the allowed. So, alpha belong to a set let us assume B which is a sub set of R p.

So, B is a p dimensional set within which alpha lies, for every point in b the differential equation is defined, but it is important to understand how the solution varies as alpha where is in the parameter space. So, much should be much of the importance of non-linear systems, deals with the variation of solution with respect to variation of parameters. Non-linear dynamics also deals with this stability properties of solutions long term behavior of the solutions. So, if one has a thorough understanding of non-linear dynamics and process related to the stability, you will be able to then appreciate the importance of the behaviour of model solution with respect to changes in parameters that is very fundamental. Because once you change the parameter some models may change their val the behaviour drastically, that leads to there is a part of the analysis of chaotic systems. So, that is something one can do.

There also other methods for approximating the solutions, there are several methods, one is called unscented transformation, we have not have time to talk about the unscented transform. There are other there is also called particle filters, particle filters is again a form of a specialized form of Monte Carlo type estimation problem to be able to do the data assimilation. So, unscented transformation base particle filter based and then there is a whole host of hybrid methodology, hybrid method that can be developed. So, once you know several algorithms you can try to combine some of the better features, the varies algorithm to create a newer algorithm for estimation, for data assimilation they are called hybridized algorithms. So, these are some of the areas one could specialize in.

There is lot of literature and non-linear dynamics, there are lot literatures in approximation of estimate in the context of non-linear models, then you do data assimilation. So, these have large potential for developing newer results and newer masters p h d thesis and potential publication.

In this course we have confined ourselves only to broad introduction to various tools and techniques for assimilating data in dynamic and static models. I am available for further help, anybody who is interested in contacting me may contact me through the email address Varahan at o u dot e d u. Varahan you have to spell it in lowercase case of course,, this is in lowercase, I hope you find these course useful and I will be I will be very happy to interact with anyone of you who want to pursue this course online, and interact with me with your questions both of the development as well as on the problems that are given to be solved. I hope you all enjoy and reap the benefits of reading through several of the reading and working through the several parts of lectures.

Thank you for this opportunity bye.