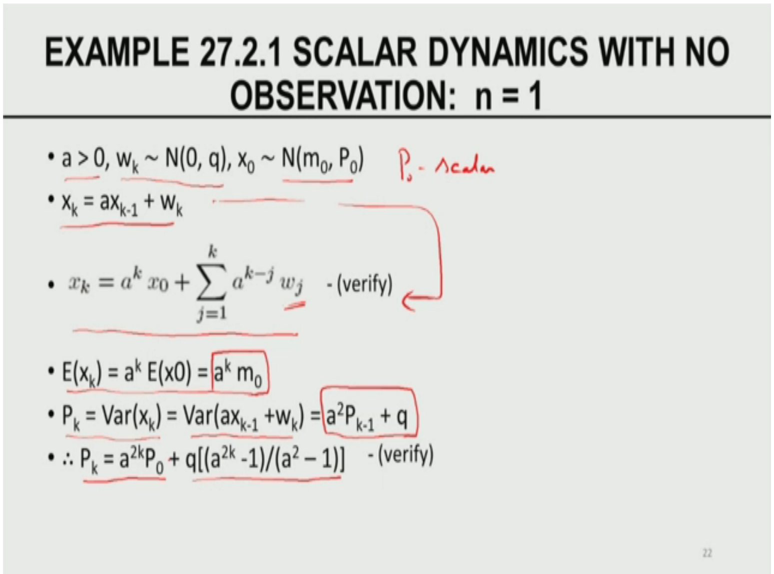
Dynamic Data Assimilation
Prof. S. Lakshmivarahan
Department of Mathematics
Indian Institute of Technology, Madras

Lecture – 34
Linear Stochastic Dynamics – Kalman Filter Continued

Now, that we have derived the equations for Kalman filter namely the forecast equation the forecast mean the forecast covariance observation the analysis step the analysis mean and analysis covariance in this case the system is linear the noise is Gaussian therefore, the forecast is a Gaussian random variable.

The observations have Gaussian distribution analysis have a Gaussian distribution as we had observed several times in the previous lectures that Gaussian distribution is the only one that is decided uniquely by the mean of the covariance. So, if I compute the analysis covariance on the analysis mean I essentially characterize the entire probability distribution against this now I am going to describe couple of simple example to illustrate the dynamics.

(Refer Slide Time: 01:30)



First I am going to talk about scalar dynamics with no observation; that means, this is called stochastic dynamics. So, let a be a scalar positive w k be a scalar Gaussian random variable with mean 0 and variance q, q is the variance and x naught is m naught P naught. So, P naught in this case is a scalar, P naught is a scalar, the dynamics is given by a simple scalar linear dynamics x k is equal to a times x k minus 1 plus w k.

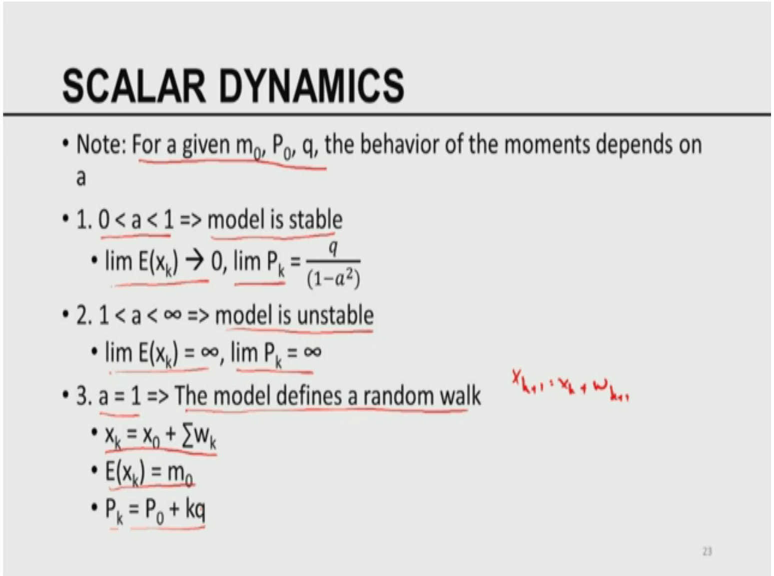
So, we have talked about the value of a w k x naught. The solution for this linear occurrence can be given by this. I would like you to verify the correctness of this equation by substituting back into the, by using the method of substitution that is the best way to describe it. So, from here we get this solution. Now I can take the mean of both sides. Please remember the mean of w j is 0 and w js are temporally uncorrelated. Therefore, if I took the mean the second term does not contribute anything. The mean of the state at time k is a to the power of k m naught the variance of this state x k at time k is P k; that is equal to variance of x a times x k minus 1 plus w k.

If you multiply a random variable by a constant you multiply the covariance by the square of the constant variance of the sum, is the sum of the variances, since the two quantities are not correlated. Therefore, the variance of the state at time k is given by this and this x k is essentially the forecast, because I am simply using the model.

So, it a to the power of k m k is the mean of the model forecast the a square P k minus 1 plus q is the variance associated with the model forecasts. Please understand this variance has two components; one coming from the initial covariance the distribution of the initial condition. Second coming from the model noise, this is the scalar analog of the vector forecasts covariance; we have already derived within the kernels within the Kalman framework.

And if you now substitute P k minus 1 in terms of P k minus 2 P k minus 2 in terms of P k minus 3 and so on, and open it up, and simplify P k depends on P naught a to the power of 2 k P naught plus q times a to the power of 2 a to the power 2 k minus 1 divided by a square minus 1. Again I would like you to verify by solving this simple linear recurrence relation.

(Refer Slide Time: 04:41)



So, for a given m naught P naught and q, what is m naught, the mean of the initial condition P naught is the covariance of the initial condition, q is the variance of the observation. I am sorry ma i is the variance of the model noise. I would like to be able to now analyze the behavior of all the moments, what are the moments the first moment and the second moment of the forecast, that simply depends on a. Now that simply depends on a, because all the other factors are fixed for a given m naught P naught and q, the behavior of the forecast moments, the first moment, second moment depends only on a.

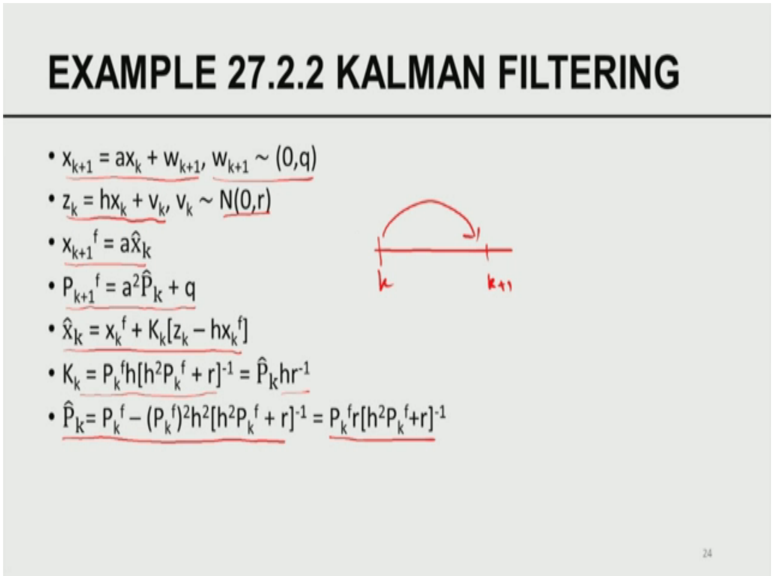
When a is less than or equal within the region a greater than 0 less than or equal to 1, the model is stable. What you mean by the model is stable. The model solution without the model solution does not explored to infinity. In fact, it can be shown from the solution of the model equation. In the previous step x k is equal to a to the power of x naught plus that we can readily verify that limit x k, as k constitutes is 0, then P k, when the limit of P k is also given by that.

So, the limit of x k is given by this, the limit of P k is given by this. And I would like to be able to tell you that this essentially comes from the first equation I, which I would like to call it star, which I would like to call it star. So, if this star, if this is double star in the next equation, both of this comes from star and double star.

Now, if a is less than, if a is less than finite, but greater than 1, the model is unstable that the complementary part the model is unstable. The limit of x k goes to infinity, the limit of P k is also goes to infinity, then a is equal to 1 the model defines the random walk. So, in this case x k plus 1 is equal to x k plus w k plus 1.

So, it executes a random walk on the real line. In this case x k is equal to x naught plus w k w k is the sum of all the noise. So, if x k is m naught P k the variance of x k P k is equal to P naught plus k q. Now you can see, even when a is equal to 1, while the beam remains the same, its covariance increases linearly as P naught plus k times q. So, this is simply the analysis of the behavior of the solution of the stochastic linear dynamics given by x k plus 1 is equal to a times x k plus w k plus 1.

(Refer Slide Time: 07:58)



Now, I would like to bring in the data into the picture, and that brings I continue the same example I am now going to talk about Kalman filtering. So, without data I simply make predictions with model alone. We talked about forecast, mean forecast covariance. we analyze how the forecast covariance varies for different regimes, when a is less than, less a is positive in between 0 and 1, when a is positive and greater than 1, when a is equal to 1.

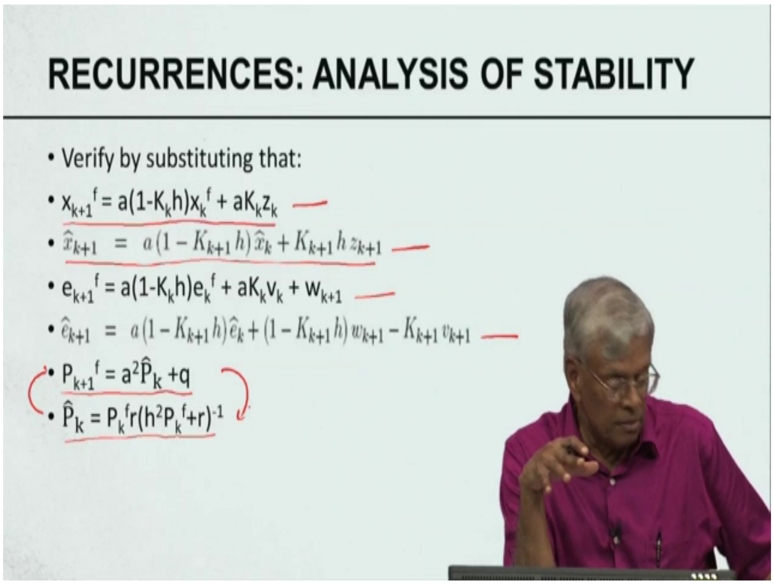
So, we divide it the range of values the parameters into three sub regions; one stable and unstable. Another corresponds to the random walk in two of the three cases when a greater than 1 or a is equal to 1, we see that the variance go to infinity, the random model is very interesting, and we may have occasion to talk about it later X k.

So, now we let us continue X k plus 1 is equal to X, w k plus 1 w k plus 1 is again the noise with 0 mean and variance, q z k is equal to h times x k plus V k h is a scalar V k is a scalar V k is 0 mean is a Gaussian random variable with 0 mean, and variance little, or now I have the distinguish between forecast and the analysis. Therefore, forecast state is equal to a times, the analysis of time k.

In other words I am going from time k to time k plus 1, this transition is what we are talking about. Earlier we talked about to transition from k minus 1 to k absolutely. There is no difference except for the values of the indices. So, P k plus 1, yeah is the forecast covariance at time k plus 1 is equal to a square times. The analysis covariance of time k plus q, the analysis itself is given by the forecast plus Kalman gain time z k minus h k f.

The Kalman gain is given by this formula, which is, which can be simplified as ,which can be simplified as P k hat h r inverse and then the analysis covariance is given by this expression P k f minus something, I am trying to subtract a positive quantity from P k f. therefore, the (Refer Time: 10:57) covariant becomes less. So, P k f or. So, analysis covariance is equal to forecast covariance times r divided by h square times P k f plus inverse. So, this is the very simple expression for the analysis covariance as a function of k. All these things arise from the derivation of the Kalman filter, except that we have substituted the corresponding formula.

(Refer Slide Time: 11:09)



So, now given I have the equations for the analysis covariance forecast forecasts covariance. I can now talk about the stability of the analysis part. In order to understand the stability of the analysis part in other words what do I want to find, does the analysis go to infinity, as time goes to infinity. If the analysis comes down to 0 what happens to the analysis that are, what happens to the forecast error as a function of the model parameter. These are some of the things that we would like to be able to understand to analyze the stability of the filter.

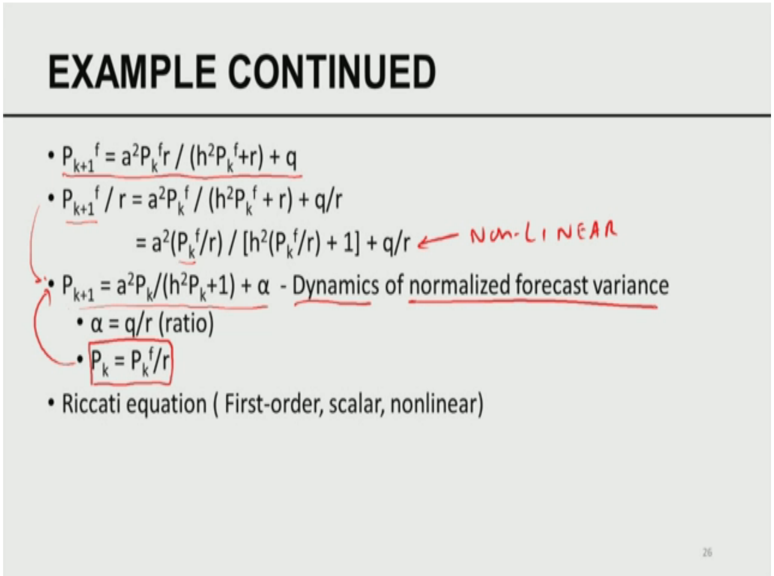
So, in this case forecast is given by this forecast 10 k plus 1 is equal to a times 1 minus k h h a K k h x k f plus a times K k z of z k the analysis. At time k is given by this. I would like you to go back to what we have, what we are doing, look at this. now I can substitute the in this equation, how do I get this. I can substitute the analysis expression into the forecast expression; that is exact. I can also substitute the forecast expression into the analysis expression.

So, that analysis at time k plus 1 can be expressed in terms of a forecast, at time k/ I am sorry analysis of time k; that means, I can get a recurrence in the analysis value at time k and k plus 1. All I can also get the forecast expressions, connecting the forecast of time k plus 1 to forecast a time k; that is the important part of the recurrence in here, analysis depends on forecast, forecast depends on analysis by mutually substituting each other. I can express forecast depend on, forecast analysis depend on analysis at time k plus 1 to time k; that is the recurrence we are talking about.

So, by substituting this, I get one recurrence for the forecast the one recurrence for the analysis. So, if I have a forecast, I can compute the forecast error, if I have analysis I can compute the analysis error. We have already seen the forecast is equal to forecast minus the state given by the model analysis. Again analysis error can again be computed as we have already seen how to handle the analysis error forecast errors, I also have the expression for the forecast covariance in terms of analysis covariance, and this case covariance is essentially a variance.

The forecast variance is the a square times the analysis variance plus q here. Again the analysis variance depends on the forecast variance. Again I can substitute each other, I can substitute this in here, I can substitute this in here, I can then relate forecast at k plus 1 to forecast at time k analysis time k plus 1 to analysis at time k, if you do that the example continues.

(Refer Slide Time: 14:38)



Now, after that substitution, I get this equation. Yes you can see this is heavily, there is a lot of heavy algebra. Now divide both sides by r. If I divide both sides by r I get this relation. So, what does it tell you. This tells you the forecast covariance are time k plus 1 is related to the forecast covariance of time k.

Now look at the expression on the right hand side, the forecast covariance occurs both in the numerator and the denominator. So, this is the non-linear recurrence relation, there is a non-linear recurrence relation.

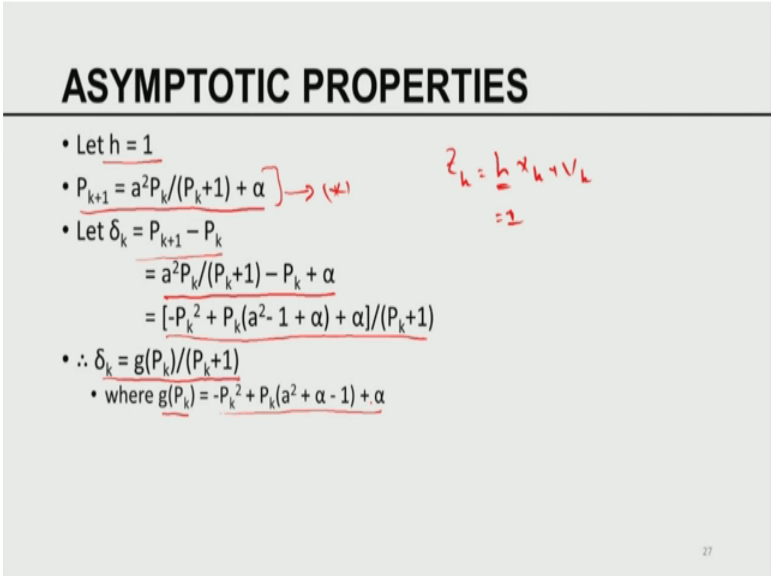
Therefore, I am now going to change the notation, I am going to define P k is equal to P k f by r. If I did this, this one in view of this one, you get this relation, where alpha is the ratio of q over r. This is an interesting ratio, what is q. q is the variance of the model noise r is the ra variance of the observation noise.

So, if the ratio of the two noise is alpha. So, if you look at this normalized forecast covariance, see that is what I want to emphasize in. This is the dynamics of evolution of their normalized forecast variance, what is the normalization. I have normalized this with respect to the variance of the observation, I have normalized this is the variance of the observation. So, if you look at this relation, you can see a is a constant P k depend P k plus 1 depends on P k h is a again a constant alpha is the ratio.

So, if dip this expression on the right hand side, depends only on a h and alpha. This type of recurrence relation in mathematics in the theory of difference equation has come to be called riccati equation. It is a first order equation, why this is the first order equation k plus 1 depends on k. Its scalar, because we are concerned with only with a radiant variances is non-linear; obviously, because the right hand side depends on P k in the both, in the numerator, as well as the denominator. Therefore, it is the first order non-linear scale of occurrence.

The particular structure has been around for a long time, it is due to riccati an Italian mathematician. There is a Riccati equation both in ordinary differential equation as well in difference equation. Here we are concerned with the difference equation, and analog of the Riccati equation. This equation is not easy to solve, because its non-linear.

(Refer Slide Time: 17:28)



So, now I am going to talk the asymptotic properties of the forecast covariance. So, let us see how you got here, once more I substituted the quantities, because forecast depends on analysis, analysis depends on forecast, I mutually substituted them and made forecast depend on forecast, analysis depend on analysis.

Likewise we did for the covariances as well. Once we did the covariances I am now trying to single out the forecast covariance. I normalized the forecast covariance by the observational covariance r, and rewrote the equation that resulted in a normalized forecast covariance dynamics, which is difference equation. It is the first order scalar non-linear difference equation.

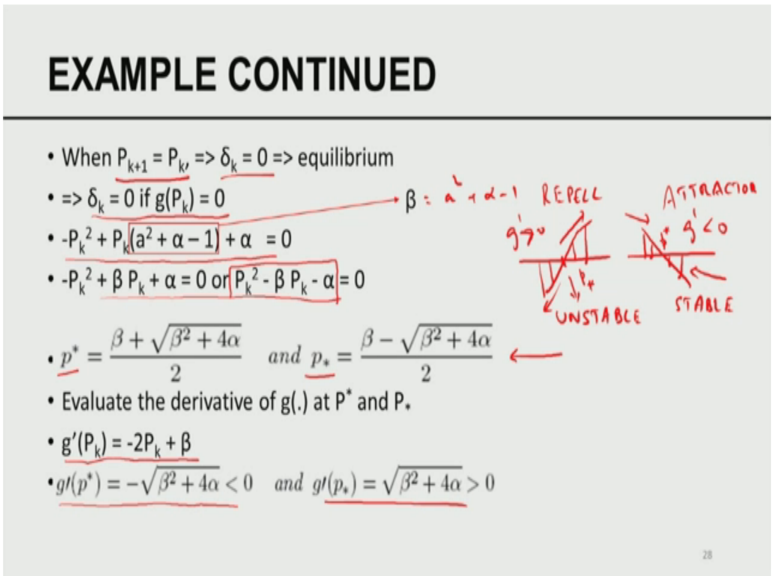
Now, I am going to look at the asymptotic properties of this first order non-linear scalar equation, why, what is the aim. The aim is the following. I would like to be able to understand the regimes, excuse me namely does the under what condition does the covariance grow, under what condition the covariance the forecast covariance died out, that relates to an analysis of stability of the forecast covariance.

Now once you analyze the stability of the forecast covariance, the corresponding properties of the analysis covariance follow immediately, because forecast covariance depends on analysis of covariance analysis, covariance depends on forecast covariance. So, behavior of one, asymptotic behavior of one will imply the asymptotic behavior of the other.

So, to that the end, I am now going to assume h is 1 without loss of generality, what is h, h is just a parameter that relates to converting the state into the observation. So, please remember z k is equal to h of x k plus v k. In this case I am assuming h is equal to 1, I am assuming h is equal to 1. So, in that case my equation now becomes simpler, like this a square P k over divided by P k plus 1 plus alpha. I would like to be able to understand the behavior, the long term behavior of this equation, this important the equation star here.

In order to understand the long-term be here of this what do I do. I am trying to look at that increment P k plus 1 minus P k; that means how much P k plus 1 differs from P k at the gate stop. If I substituted P k plus 1 in terms of P k and did the algebra. I get this relation, I get this relation, which can be now written as a ratio of a polynomial divided by P k plus 1; the polynomial P bin P k is given by the expression on the numerator, which we can readily identify. Yes, if you are trying to read through, there is a ton of algebra, and I think the only way to be able to do it, is to be able to hit all the major developments, and that is what I am trying to do leaving behind the details of the derivation of algebra to the reader.

(Refer Slide Time: 20:41)



So, when does this recurrence relation on P k converges converges; that means, P k plus 1 is equal to P k is the condition for convergence P k plus 1,1 is equal to P k is the condition for convergence at which time lambda delta k is 0, at which time delta k is 0, delta k is the equilibrium if delta k is 0. Then g of P k must be 0.

So, now, you can see how we have ah, changed the variable from P k to delta k express delta k as a ratio of two polynomials in the normalized forecast covariance P k. We have already now identified delta k being 0 is an equilibrium point, at which case P k plus 1 is equal to P k; that means, P k does not change, it is come to a stable value, because delta k is the ratio of two polynomials in P k the numerator polynomial must be 0 for delta k to 0, and that is where we are.

So, this calls for analyzing the behavior of the solution of the numerator polynomial. The numerator polynomial is equated to 0, it is rewritten, it can be rewritten by changing the variables. So, I am going to concoct a new variable beta, beta is equal to a square plus alpha minus 1. Please understand a is the model parameter alpha is the ratio of the two variances model noise to observation noise.

So, I can concoct a new symbol beta to this term. So, if I did that, my polynomial becomes this. This is second order polynomial in P k. I can now apply the standard rule for finding the roots of the second order polynomial, there are two roots P superstar, and P sub star. These expressions are now dependent on only beta, and alpha please recall beta depends on a and alpha.

So, I have a numerator polynomial g of alpha, which is a quadratic I have solved it. So, I now know the value of the equilibrium at which point, the forecast covariance will settle down. There are two equilibrium P superstar and P sub star. Now we would like to be able to understand the behavior of the solution around these two equilibria, to see whether the slope of it is increasing or decreasing; such as it is like this or if you like this, this corresponds to,, this corresponds to unstable, this corresponds to the stable.

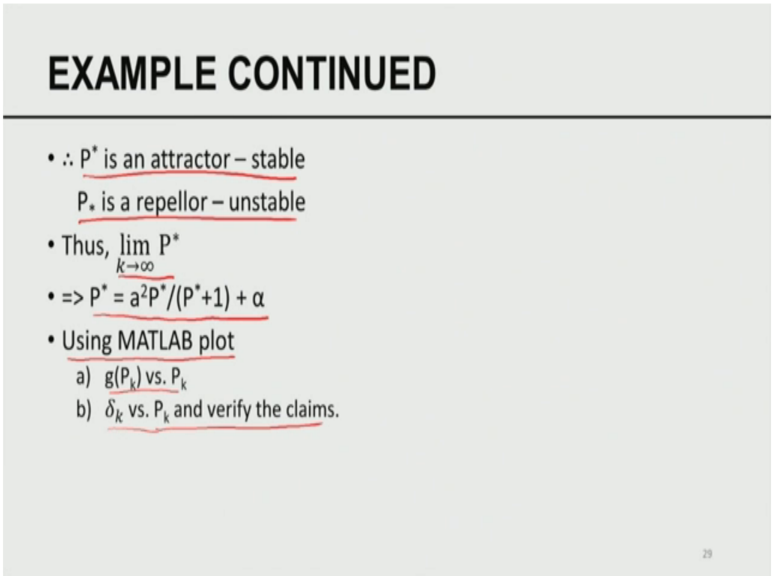
Why this corresponds to stable if I am here, this is this is P k P k plus 1 is smaller, if I am here this is P k. So, from here it pushes here from here, it pushes here; however, if I am here the. In this case if I am here, it grows bigger, if I am here it grows bigger. So, it goes away, and it goes away.

Therefore, this is unstable; that is stable, stable means if I am to the left of it, it pushes to the right, if I am to the right of it, it pushes to the left; that means, the stable equilibrium is an attractor in the neighborhood and unstable equilibrium is the repeller. If I am to the right, I move to the right. If I am to the left, I move to the left. So, this is the repel here attractor attractor.

So, to be able to see something is a repeller on attractor, we have to get the slope of g of alpha. I am sorry g of P k in the, at the point where the equilibrium occurs, where the equilibrium occurs. Therefore, the general expression for the gradient of g g prime of P k is minus 2 k minus 2 P k plus beta.

Now, we are going to evaluate by setting P k is equal to P superstar. Superstar is of the star of the superscript in this case the derivative is negative. So, please understand here the derivative is negative, here the derivative is positive. So, the derivative negative corresponds to an attractor, the derivative positive corresponds to repeller. So, you can readily identify, this is P star, this is P sub star.

(Refer Slide Time: 25:53)



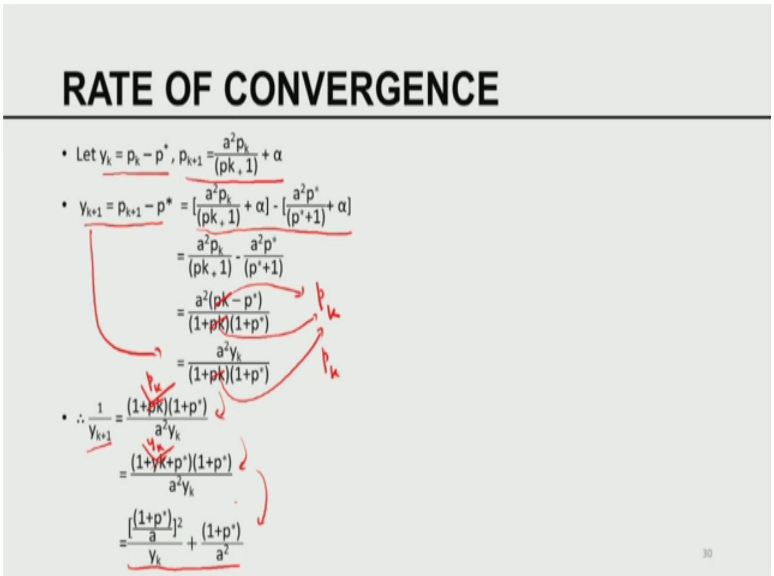
So, P sub star is unstable P star is stable. So, we need to consider only. Again I want to reemphasize one, is an attract one is attractor, another is a repeller stable unstable. Therefore, in the limit P k will go into the attractor, when P k goes into the attractor at the attractor P star is equal to a star a, a square P star divided by 1 plus P star plus alpha. So, you can readily see this must be the expression for the forecast curve, a normalized forecast covariance at the point at the equilibrium.

At the equilibrium you can solve this equation for P star, and you can get the exact value. I would like to now give strongly recommend that you plot g of P k versus P k. You know g of P k is a very simple expression, and also would like you to plot delta k versus P k, and verify all the claims that we have done. Thus for I think these are very important exercises to understand thoroughly, the long term behavior of the forecast covariance, when the model is scalar with the model is scalar.

Now, you can see I have the P star, and P superstar, and P sub start depend on beta, beta and alpha, beta depends on a. So, you can find the regions in the parameter space that gives rise to stable behaviors that gives rise to unstable behavior. I hope that is very clear. So, what does this tell you? Again I want to reemphasize this, this to essentially tells you that no matter, where you start P k plus 1 will come and settle, at this point depending on the values, depending on the values of a alpha and so on.

So, using MATLAB I would like you to be able to verify all these conclusions, and that is an important part, that is an important part of the analysis, that is the important part of a, understanding of the Kalman filter dynamics.

(Refer Slide Time: 28:26)



Once I know it converges, I think it makes sense to ask ourself the question, how fast does it converge, that relates to the rate of convergence. This is not the first time we are talking about rate of convergence; we are talked about rate of convergence, when we talked about gradient methods, and especially when iterating methods like gradient method.

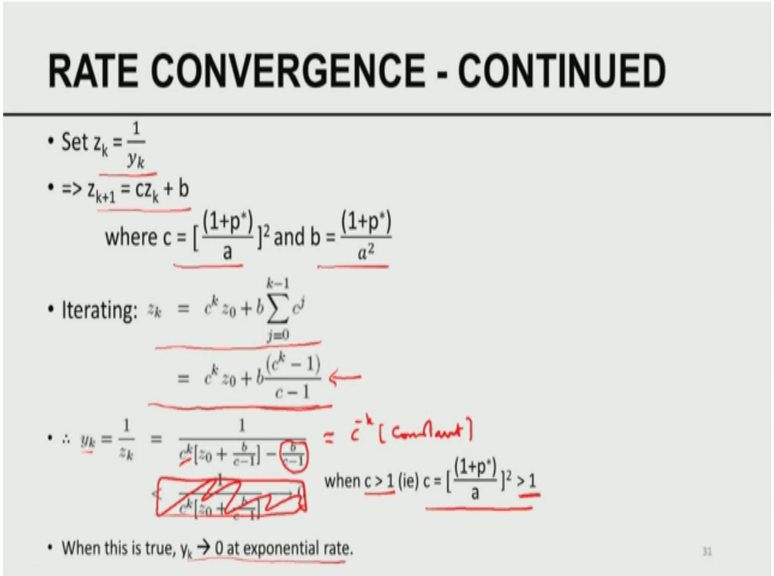
So, in any iterative algorithm, there are two questions; one should ask, does it converge if it does, at what rate. So, we are going to quickly indulge in this, in the calculation for the rate of convergence let y k vehicle to P k minus P star. So, what is P k minus P star P k is the current value of the normalized forecast, covariance P star is the asymptotic value at the stable equilibrium. So, I would like to be able to measure the difference between, where I am and where I will hit sooner or later.

So, I also know the equation for P k plus 1. So, we would like to be able to compute the rate at which, we would like to be able to compute the rate at which we converge. So, y k plus 1 is equal to P k plus 1 minus P star. So, I am going to substitute P k plus 1 P star simplify use the relation y k, this is, this must be P k. sorry P sub k P k minus P star is y k, this is also P k. Sorry this is also P k P sub k P sub k. So, if you, this is also P sub ks. So, if you can now see we have already got an expression for y k plus 1, relating to y k, and what is y k. y k is essentially the difference between P k and P star

So, if I consider 1 over y k plus 1; that is given by this expression. So, this expression is essentially coming from here, which I can rewrite like this. Again this is y k. Sorry y sub ks. Now I can rewrite this expression as a sum of these two terms. Again a little algebra will give you. So, 1 over y k plus 1 is equal to 1 over y k times a constant plus another constant. So, I am getting a recurrence for 1 over y k plus 1.

So, please go through the algebra. Now y k is the distance between P k and P star I am trying to express y k plus 1, which is the distance between P k plus 1, and P star in terms of y y k. So, I am trying to get a recurrence in y k, instead of trying to get a recurrence. I can equivalently get a recurrence in run over you y k. Why we would like to be able to consider a quantity, which is easy to analyze that is all what is the matter here is.

(Refer Slide Time: 31:44)



Now, I am going to change the variable once more. You can see how many different ways in which, you can look at it. So, 1 over y k is z k that essentially tells you the previous equation at the bottom of slide 30. Now becomes a linear equation, the linear equation is given by z k plus 1 is equal to c times z k plus b, where c is a constant and b is a constant that is the case.

So, let us go over this quickly once more, I have a riccati equation, I am going to change the variable for the riccati equation. I have a riccati equation for P k. Now I have a riccati equation, I have a corresponding non-linear equation for y k plus 1, while I have defined y k, I am trying to rewrite it as a recurrence in whatever y k, it turns out the riccati equation, after making this transformation.

These two transformation namely y k is equal to P k plus 1 minus P star 1 over y k is equal to z k the riccati equation becomes linear. Linear equation can be very easily solved; I can iterate the linear equation. So, this is the solution from the linear equation, which can be written like this. So, z k is given by this expression that is an important expression. So, 1 over z k is y k y k is equal to 1 over z k defined by this. So, let us not worry about this part.

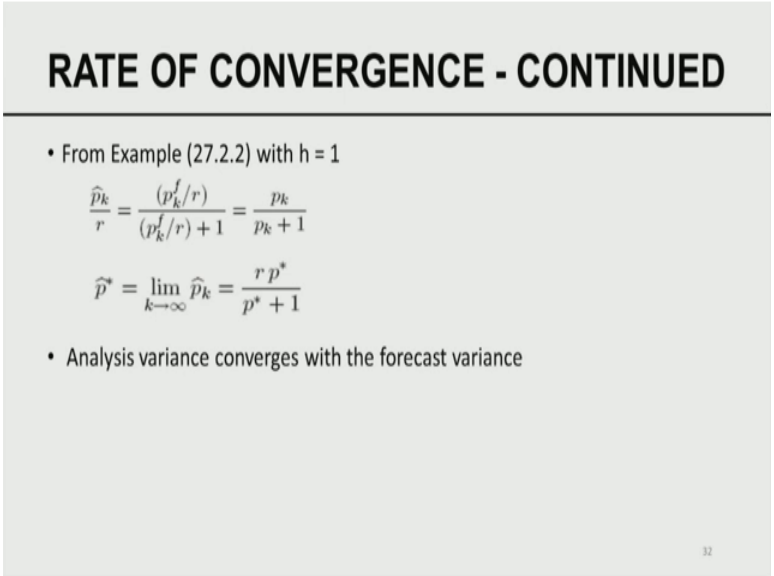
In this equation b divided by c minus 1 is a constant even. So, when c is greater than 0, and c is. I am sorry when c is greater than 1. I am sorry then c is greater than 1 c k to the power c to the power k goes to infinity. Therefore, y k will tend to 0, y k tend to 0. So, you can readily see that, this expression is equal to c to the power minus k times a constant.

Therefore when c is greater than 1, this y k tends to 0 at exponential rate. I think that is the importance of this. So, c is equal to 1 plus P star by a whole square, and that must be greater than 1. So, under the condition that c is greater than 1, which in turn relates to 1 plus P star, divided by 2 square, you are greater than 1. We can really see y k converges at an exponential rate.

So, this is an exponential convergence; that is an important; that is an important conclusion. So, what is that? We have accomplished the following by converting a sequence by converting the ricatti equation to a sequence, through a sequence of a transformation. We converted it a non-linear equation to a linear equation, which we then solved and found conditions, under which this linear equation solution to which go to infinity, which in turn means z k goes to infinity.

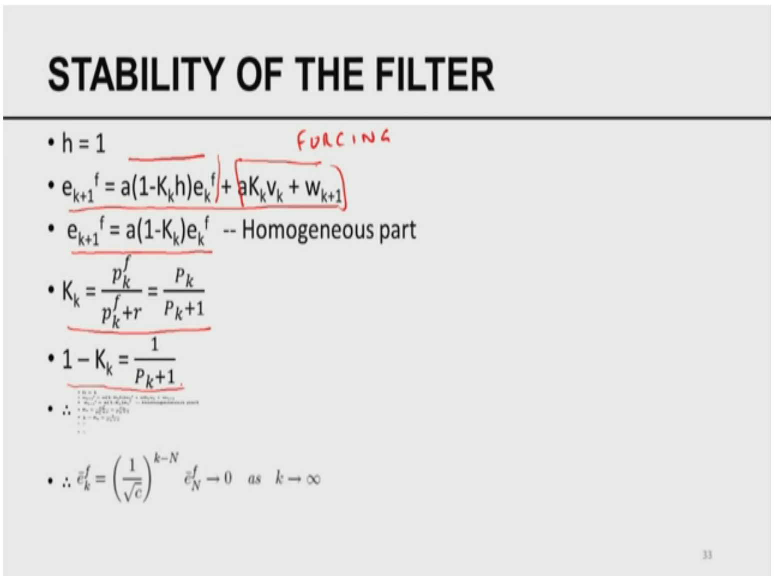
Z k goes to infinity, when c is greater than 1, when z k goes to infinity, when c greater and 1, y k tends to 0. Now please realize why the definition y k the definite y k is P k minus P star y k refers to the distance between the current value of the forecast covariance with the asymptotic value. So, that distance goes to 0 at an exponential rate, and this gives you values of the parameters under which the Kalman filter forecast covariance, not only converges, but also converge at a exponential rate. Excuse me, the rate of convergence continues.

(Refer Slide Time: 35:38)



Once we have, again we have assumed h is equal to 1. Once we have the convergence of P k, we can now conclude the convergence of P k hat, which is the analysis covariance the expression for P k hat divided by r, is given by this equation. If you took the limit of this P k hat tends to the limit, because P k f tends to a particular limit, and this is the limit of the analysis covariance. So, analysis of covariance converges with forecast covariance they work in locked step .

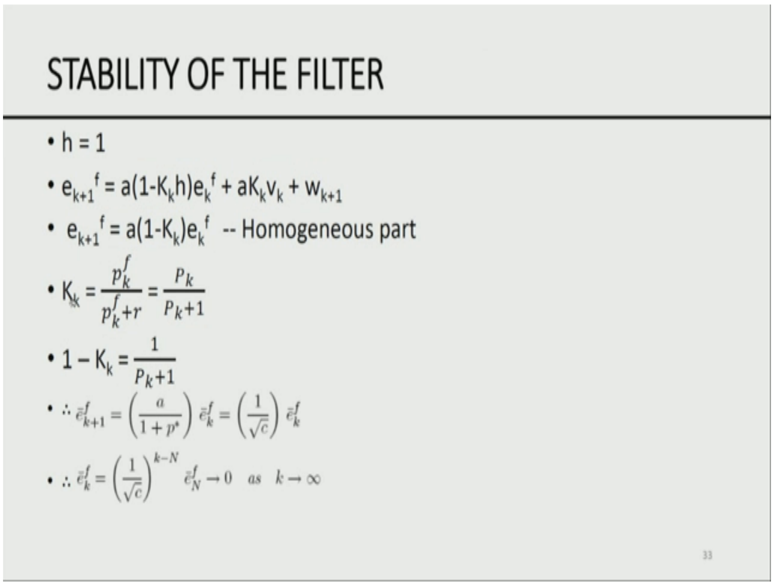
(Refer Slide Time: 36:21)



Now, I am going to go back to the analysis of what is called stability of the filter stability of the filter h is 1, this is the forecast covariance. Sorry this is the forecast covariance. I am sorry I must say forecast covariance that is wrong. This is the forecast error. The forecast error can be simplified to be this. The forecast error recurrence has two parts, the homogeneous part and the forcing part.

The homogeneous part is given at the following the forcing part, consists of two error terms a K k k v k plus w k plus 1. So, this is the stochastic part, this is the deterministic part. The deterministic part is also the homogeneous part, because this is the first thing. Please recall K k the Kalman gain is given by this. So, from here I am going to get 1 minus K k as this.

(Refer Slide Time: 37:35)



So, the Kalman gain K k is given by this formula which is P k divided P k plus 1. From that you can h readily infer 1 minus K k is equal to 1 over P k plus 1 substituting these, and simplifying it can be verified that the e bar f of k plus 1 is given by this recurrence, substituting in the value of P star. It can be verified that e bar f k plus 1 is equal to 1 over square root of c times e bar f k.

Since c is greater than 1, 1 over square root of c is less than 1; that means, in going from time k to time k plus 1, the error reduces. Therefore, if you compute the error e k bar f with respect to the state at time N e k bar f is given by 1 over square root of c to the power k minus N times e bar f by N.

So, N is, can be thought of as a starting time, since 1 over square root of N 1 over square root of c, is less than 1 as k goes to infinity. The term 1 over square root of c to the power k minus N tends to 0. Therefore, the overall error goes to 0. This in turn means the filter is stable. So, we have analyzed all the properties of the Kalman filter, and have illustrated the derivation, as well as the nuances with respect to several properties of the Kalman filter equation, using a simple static linear dynamics.

Thank you.