Signal Processing Project Work: Kalman Filtering with Measurement Losses

Sakari Teerikoski

December 23, 2014

Contents

1 Introduction

When investigating and trying to predict the outcomes of a remote phenomenon one has to either do measurements over a long distance or let measurement data from close to the interesting event travel a long distance from measurement sensors to a remote estimator. The system that is considered in this work is an estimator system containing a wireless link between sensors and a remote computer that analyzes the measurement data. There is a certain uncertainty in the receiving of measurement data over the wireless link though, and this has to be taken into account when designing a Kalman filter for predicting the state that is being measured. Basically the designed Kalman filter needs to be able to handle a situation when no measurement is received.

These situations can arise due to interference from other sources, lack of bandwith (e.g. for WLAN) or random channel variation. A loss of measurement data does not necessarily mean that a measurement is actually lost, though. It can also mean that the measurement data is delayed. But since the Kalman predictor is working with discrete time points, a delay becomes equal to a loss in this context.

2 Theory

The system that will be considered in this project has the following form:

$$x_{k+1} = Ax_k + v_k$$

$$y_k = Cx_k + w_k$$

and it is the measurement y_k that does not necessarily arrive. This uncertainty can be described as a probability λ for the measurement to arrive and a probabality $1 - \lambda$ for the opposite to happen. This defines a Bernoulli distributed random variable that will be called γ_k and its two alternating values will be set to 1 and 0 indicating the arrival and loss of measurement data respectively. To summarize this mathematically:

$$P(\gamma_k = 1) = \lambda$$

$$P(\gamma_k = 0) = 1 - \lambda$$

Due to the wireless link and the losses of measurement data one has to distinguish between the signal y_k that is sent out by the sensors and the signal that is actually received in the end. The latter signal will be called z_k here. Either the signal arrives and $z_k = y_k$, or no signal arrives so that the receiver basically receives the empty set $\emptyset$. There is no intermediate version of what is received.

In term of sequences there is one sequence $\{\gamma_k\}$ containing the packet loss information for each time step k (the receiver knows whether a signal packet is received or not) and another sequence $\{z_k\}$ containing the measurement y_k for each time step k that a measurement packet was received ($\gamma_k = 1$) and $\emptyset$, for all other time steps ($\gamma_k = 0$). To simplify notation later it is convenient to define a combined sequence $Z_k = \{\{z_k\}, \{\gamma_k\}\}$. Then simple expressions for the optimal predicted estimate and its corresponding prediction error covariance matrix is obtained:

$$\hat{x}_{k+1|k} = E[x_{k+1}|Z_k] \tag{1}$$

$$P_{k+1|k} = E[(x_{k+1} - \hat{x}_{k+1|k})(x_{k+1} - \hat{x}_{k+1|k})^T|Z_k] \tag{2}$$

The "given k" parts of the subindexed of the P-matrices will be omitted from now on. The corresponding Kalman filtering equations were obtained by Sinopoli et. al. [1]. They have the following appearence:

$$\hat{x}_{k+1|k} = A\hat{x}_{k|k-1} + \gamma_k AP_kC^T(CP_kC^T + R_2)^{-1}(y_k - C\hat{x}_{k|k-1}) \tag{3}$$

$$P_{k+1} = AP_kA^T + R_1 - \gamma_k AP_kC^T(CP_kC^T + R_2)^{-1}CP_kA^T \tag{4}$$

The initial conditions are defined as for usual Kalman filters, i.e. $\hat{x}_{0|-1} = E(x_0) = 0$ and $P_0 = R_0$.

3 Task 1: Additional theory

In the case of $\gamma_k = 1\ \forall k$ the equations (3) and (4) indeed become the Kalman filter equations in their standard predictor form [2]. However, in the case of $\gamma_k = 0$ $(\forall k)$ the result is a pair of far less accurate predictor equations. With the same initial conditions still valid, the equations become the following:

$$\hat{x}_{k+1} = A\hat{x}_k$$

$$P_{k+1} = AP_kA^T + R_1$$

This is kind of like an apriori estimate of the state. No measurement at all has been taken into account (of course since nothing arrived). This is just a naive estimator that keeps estimating the state x according to the state space model although one knows that noise giving rise to errors exsists.

In fact this is the time update (apriori estimate) formula used in the filtering form Kalman filter if one just changes $\hat{x}_{k|k-1}$ and $P_{k|k-1}$ to $\hat{x}_{k|k}$ and $P_{k|k}$. In other words this is just a Kalman predictor without its Kalman gain, so pretty much not a Kalman predictor at all.

Thus the equations in the $\gamma_k = 0$ case means exactly what $\gamma_k = 0$ would mean for a traidional Kalman predictor: it is the equation for a predictor that tries to manage without any input of measurement data (hoping that data soon will come).

From now on the system will be slightly simplified. A scalar system will be considered, where $C = 1$ and A is some scalar a with $|a| > 1$. The prediction error covariance then also becomes a scalar, denoted here as p_k. The simplified system follows the following new equations that replace the equations (3) and (4):

$$\hat{x}_{k+1|k} = a\hat{x}_{k|k-1} + \gamma_k a p_k (p_k + R_2)^{-1}(y_k - \hat{x}_{k|k-1}) \tag{5}$$

$$p_{k+1} = a^2 p_k + R_1 - \gamma_k a^2 p_k^2 (p_k + R_2)^{-1} \tag{6}$$

4 Task 2: Performance Analysis

The scalar prediction error is hard to analyse because of nonlinearities but one can find a lower and an upper bound to the expectation value $E(p_k)$. With the initial bound condition $s_0 \leq E(p_0) \leq r_0$ there will always be recursively obtainable bounds such that

$$s_k \leq E(p_k) \leq r_k. \tag{7}$$

Taking the expectation value of both sides of equation (6) gives

$$E(p_{k+1}) = E(a^2 p_k) + E(R_1) - E(\gamma_k a^2 p_k^2 (p_k + R_2)^{-1})$$

and the properties of Bernoulli distributed random variables as well as the fact that p_k clearly is independent of the same time step's γ-value gives

$$E(p_{k+1}) = a^2 E(p_k) + R_1 - a^2 \lambda E\left(\frac{p_k^2}{p_k + R_2}\right). \tag{8}$$

Because the original equation (4) is a concave function of P, one can use Jensen's inequality to find the upper bound r_k of the expectation value $E(p_k)$. Jensen's inequality states

$$E(g(p, \lambda)) \leq g(E(p), \lambda).$$

This defines an upper bound on the rational expression above:

$$E\left(\frac{p_k^2}{p_k + R_2}\right) \leq \frac{E(p_k)^2}{E(p_k) + R_2}$$

. Thus a recursive expression for the upper bound r_k has been obtained:

$$E(p_k) \leq r_{k+1} = a^2 r_k + R_1 - a^2 \lambda \left(\frac{r_k^2}{r_k + R_2}\right) \tag{9}$$

Back to equation (6) again. The bigger the negative term $\gamma_k a^2 p_k^2 (p_k + R_2)^{-1}$ is, the smaller p_{k+1} will be. Since p_{k+1} is defined to be positive semidefinite, the term will never cause the whole expression to be negative (which can be easily seen from equation (6) as well). The negative term is the biggest when R_2 is the smalles, i.e. zero, because the denominator of the term reads $p_k + R_2$. So the lowest p_{k+1} gets the expression $a^2 p_k - \gamma_k a^2 p_k + R_1$, where R_2 has been set to zero and $p_k^2 (p_k + R_2)^{-1}$ has reduced to p_k. Translating that into expectation value form as was done for the r_k case yields

$$E(p_k) \geq s_{k+1} = a^2 s_k - \lambda a^2 s_k + R_1 = (1-\lambda) a^2 s_k + R_1. \tag{10}$$

The equations (9) and (10) are the recursive expressions for the upper and lower bounds of $E(p_k)$ and the initial values for these satisfy $s_0 \leq E(p_0) \leq r_0$. This proves the claims on the boundaries given in [2].

Conditions on λ and a under which there are finite and non-negative steady state solutions to the limits $\lim_{k\to\infty} s_k = s^*$ and $\lim_{k\to\infty} r_k = r^*$ can be obtained with the help of Sinopoli's et. al. Theorem 1 [1]. It asks to consider an operator $\phi(K, X) = (1-\lambda)(AXA^T + R_1) + \lambda((A+KC)X(A+KC)^T + R_1 + KR_2K^T)$ and to suppose that $\exists \tilde{K}$ and a positive definite matrix $\tilde{P}$ such that the following is satisfied:

$$\tilde{P} > \phi(\tilde{K}, \tilde{P}).$$

Then it says that for any initial positive semidefinite matrix P_0 there exists a solution $\lim_{k\to\infty} E(P_k) = P^*$ and that P^* is the unique positive semidefinite fixed point of equation (4) above. This means that if conditions on λ and a are found for the existence of p^* they are automatically found for r^* and s^*. Sinopoli et. al. suggest $\tilde{K} = -A\tilde{P}C^T(C\tilde{P}C^T + R_1)^{-1}$ (in their Lemma 1). For the scalar system $\phi(\tilde{K}, \tilde{P})$ and $\tilde{K}$ become

$$\phi(\tilde{K}, \tilde{p}) = (1-\lambda)(a^2\tilde{p} + R_1) + \lambda((a+K)^2\tilde{p} + R_1 + KR_2K^T)$$

and

$$\tilde{K} = -\frac{a\tilde{p}}{\tilde{p} + R_1},$$

leading in the end to the following condition that has to be satisfied:

$$\tilde{p} > a^2\tilde{p} + \lambda a^2 \tilde{p}^2 (\tilde{p} + R_2)^{-2} (\tilde{p} - 2 + R_2).$$

This leads to the following conditions on λ and a for the existence of the finite non-negative steady state solution:

$$1 - \frac{1}{a^2} < \lambda \frac{\tilde{p}(2 - R_2 - \tilde{p})}{(\tilde{p} + R_2)^2} \tag{11}$$

This relation unfortunately contains the unknown $\tilde{p}$, but it can pretty much be any positive scalar.

It is known that for $\lambda = 1$ everything's fine and for $\lambda = 0$ everything fails, especially in the case of an unstable A (task 1). To prove the existence of a critical λ-value λ_c for wich and above which the expectation value's limit p^* is bounded and below which it is not, one only needs to prove that for every $\lambda > \lambda_c$ the expectation value's limit $\lim_{k\to\infty} E(p_k)$ is bounded. If there would be a $\lambda < \lambda_c$ for which that is also true, then that would contradict with the definition of λ_c. So let's compare, for $\lambda = \lambda_c$ and $\lambda = \lambda_2 \geq \lambda_c$ respectively, the expectations values $E_{\lambda_c}(p_{k+1})$ and $E_{\lambda_2}(p_{k+1})$:

$$\begin{aligned}
E_{\lambda_c}(p_{k+1}) &= E_{\lambda_c}(a^2 p_k + R_1 - \gamma_k a^2 p_k^2 (r_k + R_2)^{-1}) \\
&= a^2 E(p_k) + R_1 - a^2 \lambda_c E(p_k^2 (r_k + R_2)^{-1}) \\
&\geq a^2 E(p_k) + R_1 - a^2 \lambda_2 E(p_k^2 (r_k + R_2)^{-1}) \\
&= E_{\lambda_2}(p_{k+1}).
\end{aligned}$$

This proves that the expectation value $E_{\lambda_2}(p_{k+1})$ is lower than $E_{\lambda_c}(p_{k+1})$ for some arbitrary $\lambda_2 \geq \lambda_c$, and thus that it is bounded, which in turn proves the existence of λ_c with the following properties:

$$\lambda \geq \lambda_c \Rightarrow \lim_{k\to\infty} E(p_k) = p^* < \infty,$$

$$\lambda < \lambda_c \Rightarrow \lim_{k\to\infty} E(p_k) = \infty.$$

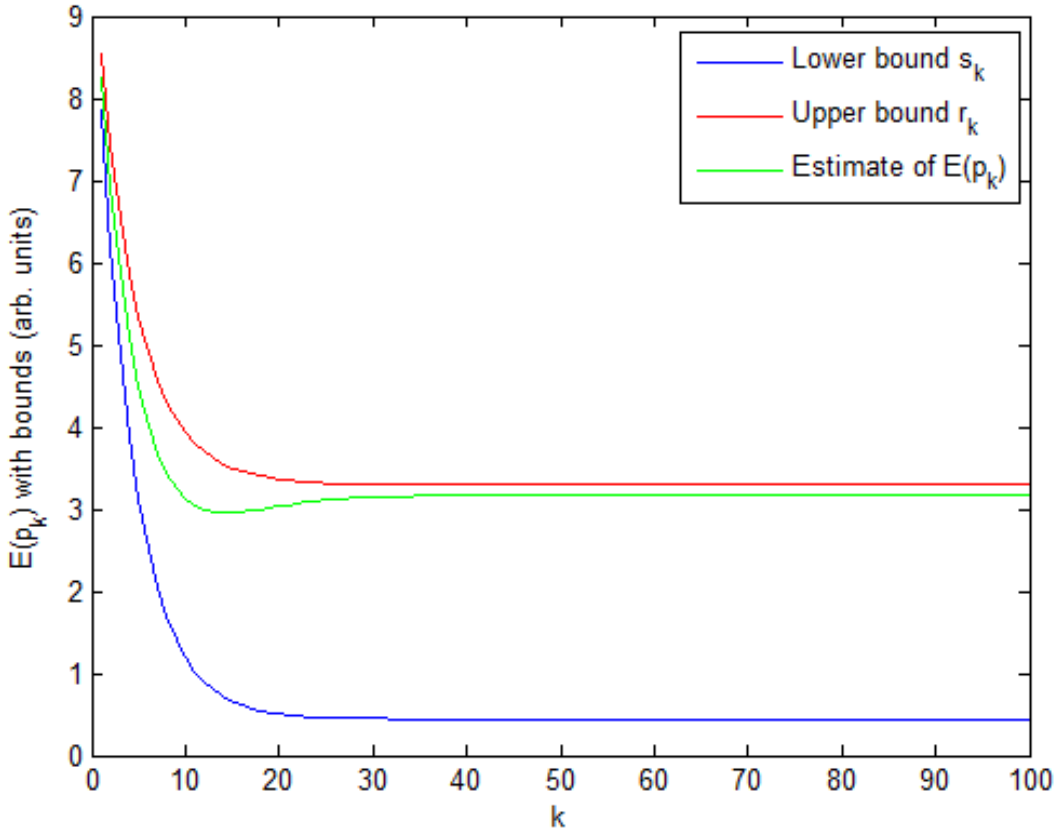

Figure 1: The stable case when $\lambda_1 = \lambda_c + 0.1$.

Assuming a scalar case where $a = 1.5$, $R_1 = 0.1$, $R_2 = 0.5$ and $E(p_0) = 10.0$, the critical probability λ_c can be calculated according to the following formula:

$$\lambda_c = 1 - 1/\alpha^2,$$

where α denotes the maximum eigenvalue of a (if a would be a matrix but in this scalar case α is a itself). This is actually the formula for the lower bound of λ_c. That will be explained later in the next section. Anyhow, λ_c becomes 0.556. The effect of the critical probability can be analyzed by calculating the aforementioned s_k and r_k for a λ_1 slightly bigger than λ_c and a λ_2 slightly less than λ_c. Graph of these have been drawn with $\lambda_1 = \lambda_c + 0.1$ and $\lambda_2 = \lambda_c - 0.1$. Figure 1 shows that both bounds converge to a stable value for $\lambda = \lambda_1$ whereas figure 2 reveales that the system is unstable for $\lambda = \lambda_2$. The s_k-values in figure 1 converge to a value of approximately 0.4 and the upper bound converges to approximately 3.3, and both values are reached around $k = 25$. In the other graph both curves diverge to positive infinity.

The expectation value $E(p_k)$ cannot be computed with an explicit expression like s_k and r_k can. To include $E(p_k)$ in the graphs one needs to estimate its value for every k. The computational problem arises from difficulties in taking the expectation value of the non-linear expression $p_k^2(p_k + R_2)^{-1}$. Just as a reminder, the expression for the expectation value is given by equation (8). This can by algebraical manipulation be written as

$$E(p_{k+1}) = a^2 E(p_k) + R_1 + a^2\lambda(E(p_k) - R_2 - R_2 E((p_k + R_2)^{-1})).$$

Now the non-linear complication has been reduced to $E((p_k + R_2)^{-1})$. In estimating $E(p_{k+1})$ one can replace $E((p_k + R_2)^{-1})$ with the average of $(s_k + R_2)^{-1}$ and $(r_k + R_2)^{-1}$ in the expression for $E(p_{k+1})$. This is a very rough approximation, though, and there are surely better ways to estimate the expectation value of p_k. The errors can be significant since the weighting of the s_k part has been put equal to the

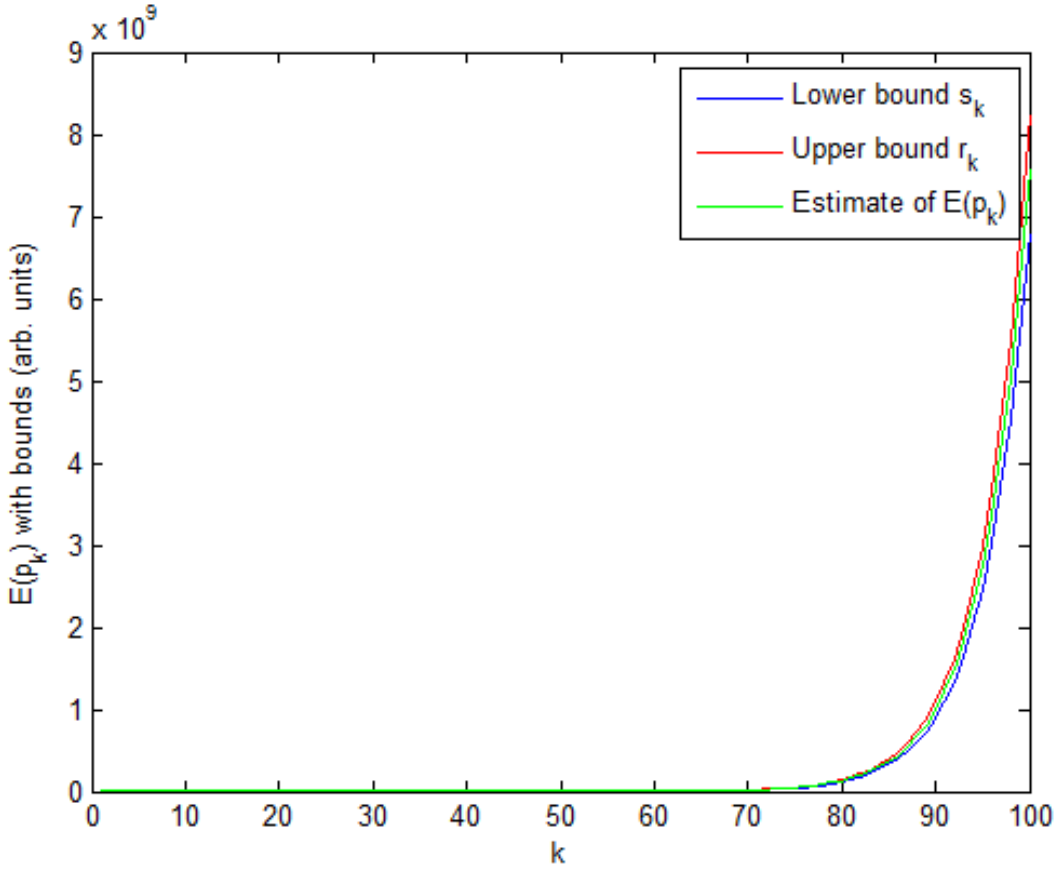

Figure 2: The diverging case when $\lambda_2 = \lambda_c - 0.1$.

weighing of the r_k part. In the graphs in figures 1 and 2, where curves for $E(p_{k+1})$ have also been included, the expectation value stays within its bounds, which is a good thing. In the stable case the area between the bound curves is quite narrow, indicating that the estimation of $E(p_{k+1})$ lies close to its real curve. However, the main problem of estimating in this way is that there is no guarantee, for the general case, that the estimation is anywhere near correct.

5 Task 3: Stability regions

Here we will consider a case with two sensors, both transmitting measurement result packets over their own wireless link. Sensor 1 takes measurements $y_k^a = x_k + v_k^a$, where v_k^a is random $N(0, \sigma_1^2)$-distributed noise. The measurement packages $\{y_k^1\}$ are lost over the first link with probability $1 - \lambda_a$. Sensor 2 takes measurements $y_k^b = x_k + v_k^b$, where v_k^b is random $N(0, \sigma_2^2)$-distributed white noise. The measurement packages $\{y_k^2\}$ are lost over the second link with probability $1 - \lambda_b$. The system equations then become

$$\begin{aligned} x_{k+1} &= ax_k + v_k \\ y_k^a &= x_k + v_k^a \\ y_k^b &= x_k + v_k^b \end{aligned}$$

.

Here it should be remembered that the C-matrix from the original state space model is $C_1 = 1$ and $C_2 = 1$ for y_k^a and y_k^b respectively. This system is obtained from an article from Liu and Goldsmith [3] and it is illustrated well in them in the block diagram displayed here in figure 3.

For λ_a and λ_b there exists a sequence of critical pairs $(\lambda_{1c}, \lambda_{2c})$ such that the expectation value of

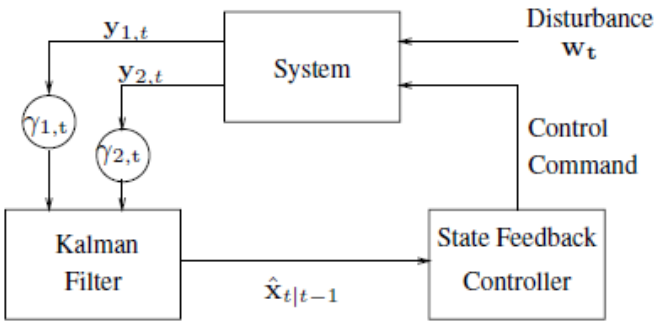

Figure 3: Block diagram for the system considered in task 3.

p_k is bounded as $k \to \infty$ if $\lambda_a > \lambda_{1c}$ and $\lambda_b > \lambda_{2c}$. Using the theory described in this section, task 3 is to approximately find this region of values for λ_a and λ_b for which $\lim_{k\to\infty} E(p_k) = \bar{p}$ such that $\bar{p}$ is finite and to draw a graph of that region for the following values: $a = 1.5$, $R_1 = 0.1$, $\sigma_1^2 = 0.2$ and $\sigma_1^2 = 0.5$.

The method for finding the values to the graph is fixing λ_1 (replacing the notations λ_a and λ_{1c} from now on) and trying to find λ_{2c} for every λ_1 from 0 to 1. A value for λ_{2c} may not always be found, but one can find its upper and lower bounds λ_{2c}^{upper} and λ_{2c}^{lower}.

The system can be described as a MIMO system:

$$x_{k+1} = ax_k + v_k$$

$$\begin{bmatrix} y_k^a \\ y_k^b \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix} x_k + \bar{w}_k$$

The noise vector combines v_k^a and v_k^b and results in a redefinition of $R2$:

$$\begin{bmatrix} \sigma_1^2 & 0 \\ 0 & \sigma_2^2 \end{bmatrix}.$$

Also there is a new C-matrix $C = \begin{bmatrix} C_1 & C_2 \end{bmatrix}^T = \begin{bmatrix} 1 & 1 \end{bmatrix}^T$. The uncertainty variables $\gamma_{1,k}$ and $\gamma_{2,k}$ in the block diagram in figure 3 share the same Bernoulli distribution properties as in the previous cases and their corresponding probabilities are the given lambdas. The corresponding equations to the equations (3) and (4) in the sace with two wireless links are

$$\hat{x}_{k+1|k+1} = \hat{x}_{k+1|k} + p_{k+1|k} \begin{bmatrix} 1 & 1 \end{bmatrix} \left(\begin{bmatrix} 1 \\ 1 \end{bmatrix} p_{k+1|k} \begin{bmatrix} 1 & 1 \end{bmatrix} + R_2 \right)^{-1} \begin{bmatrix} 1 \\ 1 \end{bmatrix} \left(\bar{y}_{k+1} - \begin{bmatrix} 1 \\ 1 \end{bmatrix} \hat{x}_{k+1|k} \right) \tag{12}$$

and

$$p_{k+1|k+1} = p_{k+1|k} - p_{k+1|k} \begin{bmatrix} 1 & 1 \end{bmatrix} \left(\begin{bmatrix} 1 \\ 1 \end{bmatrix} p_{k+1|k} \begin{bmatrix} 1 & 1 \end{bmatrix} + R_2 \right)^{-1} \begin{bmatrix} 1 \\ 1 \end{bmatrix} p_{k+1|k} \tag{13}$$

respectively. These are when $\gamma_{1,k}$ and $\gamma_{2,k}$ are always one. When they are always zero the equations (12) and (13) take less interesting forms:

$$\hat{x}_{k+1|k+1} = \hat{x}_{k+1|k},$$

$$p_{k+1|k+1} = p_{k+1|k}.$$

If one of them is always received and the other is never received, the equations become the standard Kalman filter with the received measurements as the only one, for both cases. The general case is, however, not any of these. The general expression, as derived by Liu and Goldsmith, of p_{k+1} is the following:

$$\begin{aligned} p_{k+1} = a^2 p_k + R_1 - \gamma_{1,k}\gamma_{2,k} a^2 p_k^2 \begin{bmatrix} 1 & 1 \end{bmatrix} \left(\begin{bmatrix} 1 \\ 1 \end{bmatrix} p_k \begin{bmatrix} 1 & 1 \end{bmatrix} + R_2 \right)^{-1} \begin{bmatrix} 1 \\ 1 \end{bmatrix} \\ - \gamma_{1,k}(1-\gamma_{2,k}) a^2 p_k^2 (p_k + \sigma_1^2)^{-1} - (1-\gamma_{1,k})\gamma_{2,k} a^2 p_k^2 (p_k + \sigma_2^2)^{-1}. \end{aligned} \tag{14}$$

It is said in [3] that the sequence $\{p_k\}_{k=0}^{\infty}$ is a random process given an initial condition p_0 and that p_k is bounded with probability 1 if and only if $E(p_k)$ is bounded. $E(p_{k+1})$ is equal to the expectation of what is obtained with input data p in the following so called modified algebraic Riccati equation:

$$\begin{aligned} g_{\lambda_1\lambda_2}(X) = a^2X + R_1 - \lambda_1\lambda_2 a^2X^2 \begin{bmatrix} 1 & 1 \end{bmatrix} \left(\begin{bmatrix} 1 \\ 1 \end{bmatrix} X \begin{bmatrix} 1 & 1 \end{bmatrix} \right) \begin{bmatrix} 1 \\ 1 \end{bmatrix} \\ - \lambda_1(1-\lambda_2)a^2X^2(X+\sigma_1^2)^{-1} - (1-\lambda_1)\lambda_2 a^2X^2(X+\sigma_2^2)^{-1}. \end{aligned} \tag{15}$$

Liu and Goldsmith provide useful mathematical theorem for the determination of λ_{2c}. A short summary of the most relevant ones will be given here:

If $g_{\lambda_1\lambda_2}(p_k)$ is unstable for $\lambda_b = 0$ and stable for $\lambda_b = 0$, then the λ_{2c}-value exists. For a given λ_1, one can compute a lower and an upper bound for λ_{2c} with the formulas

$$\lambda_{2c}^{lower} = \max\left(1 - 1/\alpha^2(1-\lambda_1), 0\right),$$

and

$$\lambda_{2c}^{upper} = argin f_{\lambda_2}(\exists K, K_1, K_2, X | X > \phi(K, K_1, K_2, X)),$$

where α is the maximum eigenvalue of a, so a itself in this case.

Using the equation

$$\lambda_{2c}^{lower} = \max\left(1 - \frac{1}{\alpha^2(1-\lambda_1)}, 0\right),$$

where $\alpha = a$ for the scalar case considered here, a graph for the lower bound of the critical λ-value is finally obtained. This is shown in figure 4.

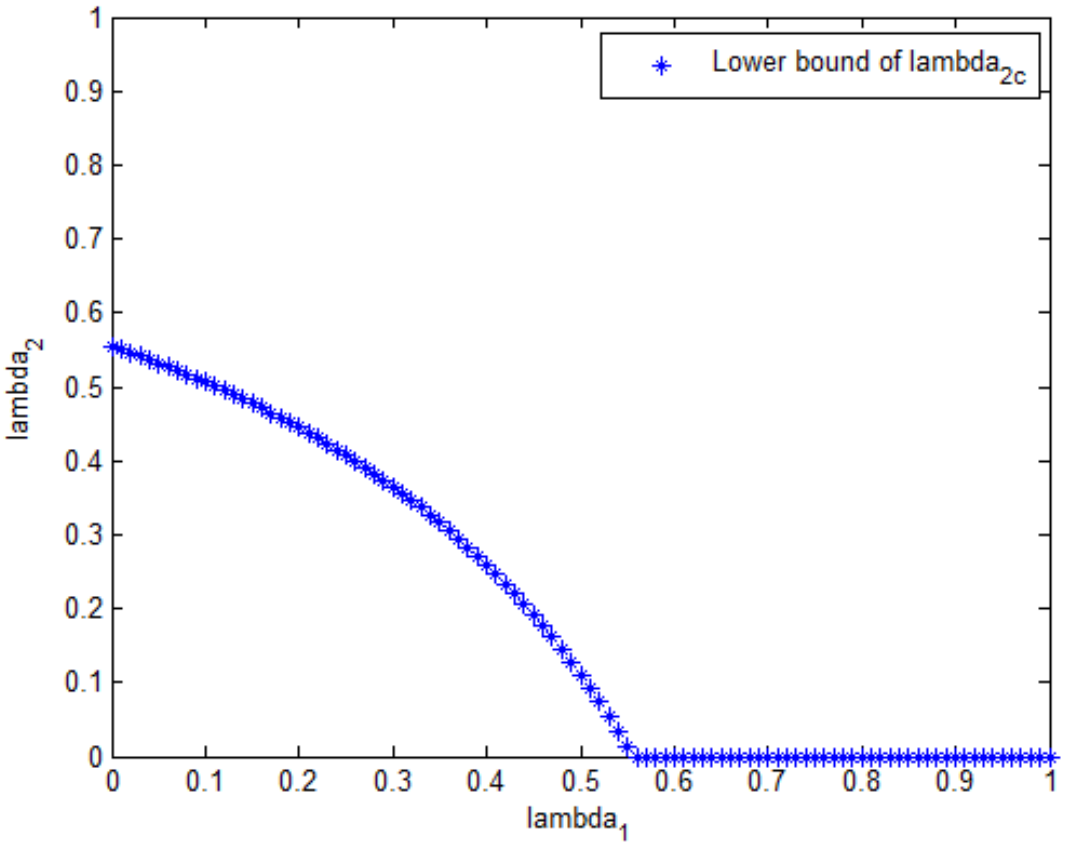

Figure 4: The stability region's lower bound.

The next step is to find the upper bound. One can intuitively think that it should show the same kind of behaviour in the graph as the lower bound. Actually, before computing the upper bound, it should be stated that a section in the work of Liu and Goldsmith [3] actually gives away the answer: "When C_1 and C_2 are both square and invertible, the lower bound λ_{2c}^{lower} and the upper bound λ_{2c}^{upper} are equal for every λ_1". In this scalar case, both C_1 and C_2 is equal to 1, so the lower and upper bound should coincide. However, let us not ruin the fun of computing stuff! The phi-function is defined according to

$$\phi(K, K_1, K_2, X)) = (1-\lambda_1)(1-\lambda_2)(a^2X+R_1)+\lambda_1\lambda_2(f^2X+V)+\lambda_1(1-\lambda_2)(f_1^2X+V_1)+\lambda_2(1-\lambda_1)(f_2^2X+V_2), \tag{16}$$

where the parameters f and V are defined in [3] as:

$$f = a + K\begin{bmatrix}1\\1\end{bmatrix},$$

$$f_1 = a + K_1$$

$$f_2 = a + K_2$$

$$V = R_1 + KR_2K^T$$

$$V_1 = R_1 + K_1^2\sigma_1^2$$

$$V_2 = R_1 + K_2^2\sigma_2^2$$

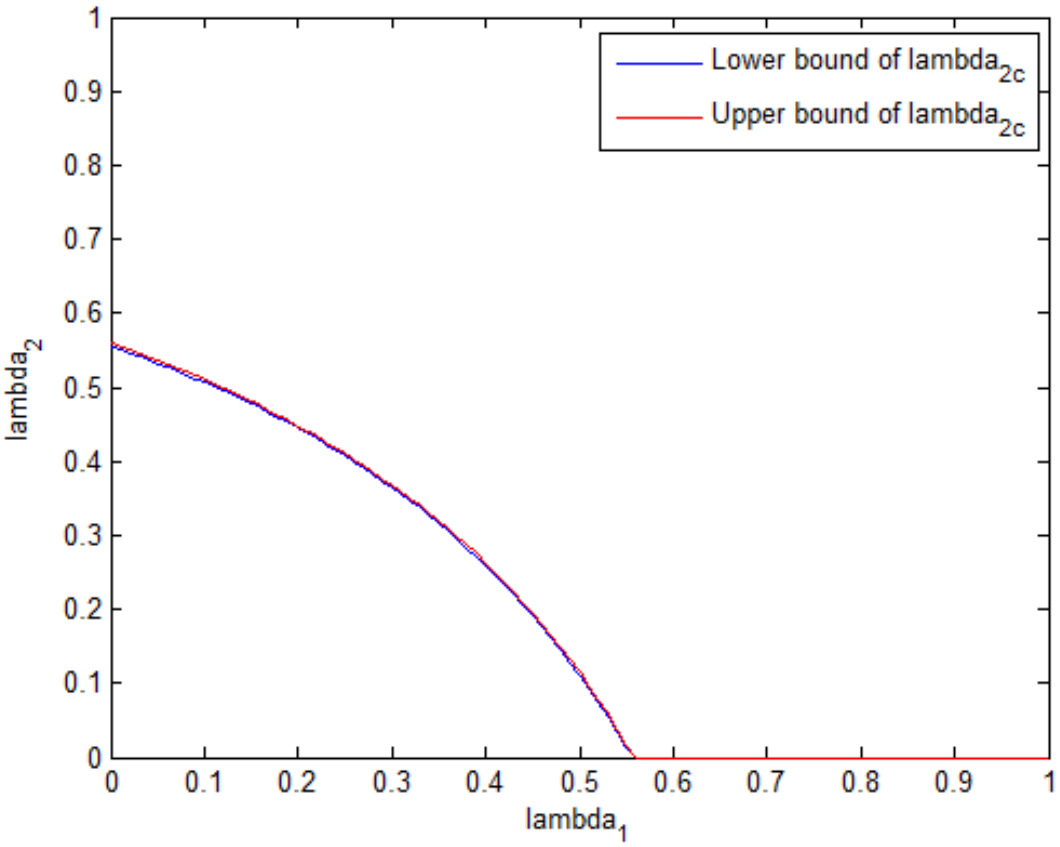

Figure 5: The stability region with upper and lower bounds displayed.

Liu and Goldsmith showed that the lowest value for $g_{\lambda_1\lambda_2}(X)$ is obtained with the X-dependent values $K = -aXC^T(CXC^T + R_2)^{-1}$, $K_1 = -aX(X + \sigma_1^2)^{-1}$ and $K_2 = -aX(X + \sigma_2^2)^{-1}$. The theorems

require $X > 0$ (because X in this function corresponds to p_k) and they give the last condidtion that has to be satisfied: $X - \phi > 0$. Since for these data the g-function is the minimum, $X - \phi > 0$ not only should be satisfied but the absolute value $X - \phi$ should also be minimized to find λ_{2c}. This has been done in the MATLAB code that is submitted with this report. It was used to solve the inequality yielding λ_{2c}^{upper} and it was used to plot all the graphs. The upper bound case is solved with solving a few inequalities in MATLAB. In the end we get a graph with both of the bounds as shown in figure 5. The stability region is then the region in (λ_a, λ_b)-space than lies above the curve in figure 5.

6 References

1. B. Sinopoli, L. Schenato, M. Francheschetti, K. Poolla, M. Jordan, S. Sastry, *Kalman Filtering with intermittent observations*, in *IEEE Transactions on Automatic Control*, vol. 49, no. 9, pp. 1453-1464, 2004

2. Unknown author, the assignment text to this project

3. X. Liu, A. Goldsmith, *Kalman Filtering with partial observation losses*, in *Proceedings of IEEE CDC* 2004, pp. 4180-4186, Atlantis, Paradise Island, Bahamas, 2004