I. SUMMARY

Here are the key ideas and results of this important topic.

- Section II reviews Kalman Filter.
- A system is observable if its state can be determined from its outputs (after some delay).
- A system is reachable if there are inputs to drive it to any state.
- We explore the evolution of the covariance in a linear system in Section IV.
- The error covariance of a Kalman Filter is bounded if the system is observable.
- The covariance increases if it starts from zero.
- If a system is reachable and observable, then everything converges and the stationary filter is asymptotically optimal.

II. REVIEW: KALMAN FILTER

We consider a linear system with matrices that do not depend on time. That is,

$$X_{n+1} = AX_n + V_n$$
and
$$Y_n = CX_n + W_n, n \geq 1,$$

where \{X_1, V_n, W_n, n \geq 1\} are all orthogonal and zero-mean with \text{cov}(V_n) = K_V \text{ and } \text{cov}(W_n) = K_W.

Theorem 1: Kalman Filter

\[
\hat{X}_n = L[X_n|Y_1, \ldots, Y_n] = L[X_n|Y^n]
\]

is obtained by the Kalman Filter:

\[
\hat{X}_n = A\hat{X}_{n-1} + R_n(Y_n - CA\hat{X}_{n-1})
\]

where

\[
R_n = S_nC^T [CS_nC^T + K_W]^{-1},
\]

\[
S_n = A\Sigma_{n-1}A^T + K_V,
\]

\[
\Sigma_n = (I - R_nC)S_n.
\]

Thus,

\[
S_{n+1} = K_V + AS_nA^T - AS_nC^T [CS_nC^T + K_W]^{-1}CS_nA^T.
\]

Moreover, the matrices \(S_n\) and \(\Sigma_n\) have the following significance:

\[
S_n = \text{cov}(X_n - A\hat{X}_{n-1}),
\]

\[
\Sigma_n = \text{cov}(X_n - \hat{X}_n).
\]

Proof:

Let

\[
U_n = Y_n - L[Y_n|Y^{n-1}] = Y_n - L[CX_n + W_n|Y^{n-1}] = Y_n - CL[X_n|Y^{n-1}] = Y_n - CA\hat{X}_{n-1} = CX_n + W_n - CL[X_n|Y^{n-1}]
\]

Then, by (1) in Lecture 12,

\[
\hat{X}_n = L[X_n|Y^n] = L[X_n|Y^{n-1}] + L[X_n|U_n] = A\hat{X}_{n-1} + R_nU_n = A\hat{X}_{n-1} + R_n(Y_n - CA\hat{X}_{n-1})
\]

where

\[
R_n = \text{cov}(X_n, U_n)\text{cov}(U_n)^{-1}.
\]

Now,

\[
\text{cov}(X_n, U_n) = \text{cov}(X_n, CX_n + W_n - CL[X_n|Y^{n-1}])
\]

\[
= \text{cov}(X_n - L[X_n|Y^{n-1}], CX_n + W_n - CL[X_n|Y^{n-1}]) = S_nC^T.
\]

Also,

\[
\text{cov}(U_n) = \text{cov}(CX_n + W_n - CL[X_n|Y^{n-1}]) = CS_nC^T + K_W.
\]
Thus,
\[ R_n = S_n C^T [C S_n C^T + K_W]^{-1}. \]

In addition,
\[ S_n = \text{cov}(X_n - L[X_n|Y^{n-1}]) = \text{cov}(AX_{n-1} + V_{n-1} - A\hat{X}_{n-1}) = A\Sigma_{n-1} A^T + K_V. \]

Finally, by (1) in Lecture 12,
\[ \Sigma_n = \text{cov}(X_n - L[X_n|Y^n]) = \text{cov}(X_n - L[X_n|Y^{n-1}]) - \text{cov}(X_n, U_n)\text{cov}(U_n)^{-1}\text{cov}(U_n, X_n) = S_n - R_n C S_n. \]

In the following sections, we explore what happens as \( n \to \infty \). Specifically, we want to understand if the error covariance \( \Sigma_n \) blows up or if it converges to a finite value. First, we recall some key notions about linear systems.

III. OBSERVABILITY AND REACHABILITY

**Lemma 1**: Cayley-Hamilton
Let \( A \in \mathbb{R}^{m \times m} \) and \( \det(sI - A) = \alpha_0 + \alpha_1 s + \cdots + \alpha_{m-1} s^{m-1} + s^m. \) Then
\[ \alpha_0 I + \alpha_1 A + \cdots + \alpha_{m-1} A^{m-1} + A^m = 0. \]
In particular, \( \text{span}\{ I, A, A^2, \ldots \} = \text{span}\{ I, A, \ldots, A^{m-1} \}. \)

**Definition 1**: Observability and Reachability
The linear system
\[ X_{n+1} = AX_n, Y_n = CX_n, n \geq 1 \]
is **observable** if \( X_1 \) can be determined exactly from \( \{ Y_n, n \geq 1 \} \). We then say \( (A, C) \) is observable.

The linear system
\[ X_{n+1} = AX_n + CU_n, n \geq 1 \]
is **reachable** if, for every state \( X \), there is a sequence of inputs \( \{ U_n, n \geq 1 \} \) that drives the system from \( X_1 = 0 \) to state \( X \). We say that \( (A, C) \) is reachable.

**Fact 1**: (a) \( (A, C) \) is observable if and only if \( [C^T A^T C^T] \cdots [(A^{m-1})^T C^T] \) is of full rank.
(b) \( (A, C) \) is reachable if and only if \( (A^T, C^T) \) is observable, i.e., if and only if \( [C|AC|\cdots|A^{m-1}C] \) is of full rank. In that case,
\[ \sum_{p=0}^{m-1} A^p C C^T (A^p)^T \]
is positive definite.

**Proof**: To see (a) note that (1) implies that \( Y_n = CX_n = CA^{n-1}X_1 \). Consequently, observability is equivalent to the null space of \( [C|CA|CA^2|\cdots|] \) being \( \{ 0 \} \). Accordingly, this is equivalent to the matrix \( [C^T A^T C^T] [(A^2)^T C^T] |\cdots| \) being of full rank. The conclusion then follows from Lemma 1.

For (b), observe that (2) implies that
\[ X_n = A^{n-1}X_1 + \sum_{k=1}^{n-1} A^{n-k-1}CU_k = \sum_{k=1}^{n-1} A^{n-k-1}CU_k. \]
Therefore, the system is reachable if and only if \( [C|AC|A^2C|\cdots|] \) is of full rank. The conclusion then follows again from Lemma 1.

IV. SYSTEM ASYMPOTICS

Our discussion is borrowed from [2]. First, let us examine the evolution of the unobserved system
\[ X_{n+1} = AX_n + V_n, n \geq 1, \]
where \( \{ X_1, V_n, n \geq 1 \} \) are all orthogonal and zero-mean with \( \text{cov}(V_n) = K_V \) and \( K_n = \text{cov}(X_n) \). Note that
\[ K_{n+1} = AK_n A^T + K_V. \]

The following theorem describes the evolution of \( K_n \). Recall that a matrix \( A \in \mathbb{R}^{m \times m} \) is said to be stable if its eigenvalues all have magnitude strictly less than 1.

**Theorem 2**: (a) If \( A \) is stable, then there is a positive semidefinite matrix \( K \) such that \( K_n \to K \) as \( n \to \infty \). Moreover, \( K \) is the unique solution of the equation
\[ K = AK A^T + K_V. \]
(b) Let $K_V = QQ^T$. Assume that $(A,Q)$ is reachable. Then $A$ is stable if and only if the equation

$$K = AK^T + QQ^T$$

has a positive definite solution $K$.

**Proof:** (a) By induction we see that the solution of (3) is

$$K_n = A^{n-1}K_1(A^{n-1})^T + \sum_{p=0}^{n-2} A^p K_V (A^p)^T.$$  \hspace{1cm} (5)

If $A$ is stable, then $|A^{n+1}| \leq C \alpha^n$ for some finite $C$ and some $\alpha \in (0,1)$, as it can be seen by considering the SVD of $A$. This implies that the first term in (5) vanishes and the sum converges to some $K = \sum_{p=0}^{\infty} A^p K_V (A^p)^T$ that is clearly positive semidefinite. The convergence also implies (4). It remains to show that (4) has a unique solution. Let $K'$ be another solution. Then $\Delta = K' - K$ satisfies $\Delta = A \Delta A^T$. Recursive substitutions show that $\Delta = A^n \Delta (A^n)^T$. Letting $n \to \infty$ shows that $\Delta = 0$ and the solution of (4) is unique.

(b) If $A$ is stable, we know from (a) that (4) has a unique positive semidefinite solution $K = \sum_{p=0}^{\infty} A^p QQ^T (A^p)^T$. Since $(A,Q)$ is reachable, the null space of $[Q^T|Q^TA^T|Q^T(A^2)^T|\ldots]$ is $\{0\}$. This implies $K$ is positive definite.

To show the converse, assume that $K$ is a positive definite solution of (4). Then we know that $K = A^n K (A^n)^T + \sum_{p=0}^{n-1} A^p QQ^T (A^p)^T$, $n \geq 1$. Let $\lambda$ be an eigenvalue of $A^T$ with eigenvector $v$. Then $A^T v = \lambda v$ and

$$v^* K v = |\lambda|^{2n} v^* K v + v^* \left( \sum_{p=0}^{n-1} A^p QQ^T (A^p)^T \right) v.$$ 

But $\sum_{p=0}^{n-1} A^p QQ^T (A^p)^T$ is positive definite from the reachability of $(A,Q)$, which implies that the last term in the above identity is positive. Consequently, it must be that $|\lambda| < 1$.

The statements of the theorem should be intuitive. If the system is stable, then the state tries to go to zero but is constantly pushed by the noise. That noise cannot push the state very far and one can expect the variance of the state to remain bounded. The convergence is a little bit more subtle. If the system is reachable, then the noise pushes the state in all directions and it is not surprising that the variance of the state is positive definite if the system is stable. If the system is not stable, the variance would explode.

V. FILTER ASYMPTOTICS

We now explore the evolution of Kalman Filter.

**Theorem 3:** Let $K_V = QQ^T$. Suppose that $(A,Q)$ is reachable and $(A,C)$ is observable. If $S_1 = 0$, then

$$\Sigma_n \to \Sigma, R_n \to R, S_n \to S, \text{ as } n \to \infty.$$ 

The limiting matrices are the only solutions of the equations

$$\Sigma = (I - RC) S, R = SC^T (CSC^T + K_V)^{-1}, \quad S = A \Sigma A^T + K_V.$$ 

Equivalently, $S$ is the unique positive semidefinite solution of

$$S = A (S - SC^T (CSC^T + K_V)^{-1}CS) A^T + K_V.$$  \hspace{1cm} (6)

Moreover, the time-invariant filter $Z_n = A Z_{n-1} + R(Y_n - CA Z_{n-1})$ satisfies $\text{cov}(Z_n - \hat{Z}_n) \to \Sigma$, as $n \to \infty$.

**Comments:** The reachability implies that the noise excites all the components of the state. The observability condition guarantees that the observations track all the components of the state and imply that the estimation error remains bounded. Note that the state could grow unbounded, if $A$ is unstable, but the estimator tracks it even in that case. The time-invariant filter has the same asymptotic error as the time-varying one.

**Proof:** The proof has the following steps. For two positive semidefinite matrices $S$ and $S'$, we say that $S \leq S'$ if $S' - S$ is positive semidefinite. Similarly, we say that the positive semidefinite matrices $\{S_n, n \geq 1\}$ are bounded if $S_n \leq S$ for some positive semidefinite matrix $S$.

• (a) The matrices $S_n$ are bounded.
• (b) If $S_1 = 0$, then $S_n$ is nondecreasing in $S_1$.
• (c) If $S_1 = 0$, then $S_n \uparrow S$, where $S$ is a positive semidefinite matrix.
• (d) The matrix $A - ARC$ is stable.
• (e) For any $S_1, S_n \to S$.
• (f) Equation (6) has a unique positive semidefinite solution $S$.
• (g) The time-invariant filter has the same asymptotic error covariance as the time-varying filter.

We outline these steps.
(a) The idea is that, because of observability, $X_{n+m}$ is a linear function of $\{Y_p, V_p, W_p, p = n, \ldots, n + m - 1\}$. The covariance $S_{n+m}$ must be bounded by that of $X_{n+m}$ given $\{Y_p, p = n, \ldots, n + m - 1\}$, which is a linear combination of the covariances of $2m$ random variables and is therefore uniformly bounded for all $n$.

(b) That is, if $S_1$ is replaced by $S'_1 \ge S_1$, then $S_n$ is replaced by $S'_n \ge S_n$. The proof of this fact is pretty neat. One could try to do it by induction, based on the algebra. That turns out to be tricky. Instead, consider the following argument. Since we worry only about covariances, we may assume that all the random variables are jointly Gaussian. In that case, increasing $S_1$ to $S'_1 > S_1$ can be done by replacing $X_1$ by $X'_1 = X_1 + \xi_1$, where $\xi_1 = N(0, S'_1 - S_1)$ and is independent of $\{X_i, V_n, W_n, n \ge 0\}$. Now, let $X'_n$ be the system state corresponding to $X'_1$ and $Y'_n$ be the corresponding observation. Note that $X'_{n+1} = X_{n+1} + A^n\xi_1$. Consequently,

$$L[X'_{n+1}|Y^n, \xi_1] = L[X_{n+1}|Y^n] + L[X_{n+1}|\xi_1] + L[A^n\xi_1|Y^n] + L[A^n\xi_1|\xi_1] = L[X_{n+1}|Y^n] + A^n\xi_1$$

and

$$S_{n+1} = \text{cov}(X'_{n+1} - L[X'_{n+1}|Y^n], \xi_1)$$

while

$$S'_{n+1} = \text{cov}(X'_{n+1} - L[X'_{n+1}|(Y')^n]).$$

Since, for each $n \ge 1$, one can express $Y'_n$ as a linear function of $(Y_n, \xi_1)$, it follows that $S_{n+1} \le S'_{n+1}$.

(c) Note that $S_1 = 0 \le S_2$. From part (b), $S_n \le S_{n+1}$. However, $S_n \le B$, for some positive semidefinite $B$. Thus, $S_n(i, i) := c^2 + S_n e_i \le S_{n+1}(i, i) \le B(i, i)$. This implies that $S_n(i, i) \uparrow \gamma S(i, i)$ for some nonnegative $S(i, i)$. Similarly, $\alpha(n) := S_n(i, i) + 2S_n(i, j) + S_n(j, j) = (e_i + e_j)^T S_n(e_i + e_j) \le \alpha(n + 1) \le (e_i + e_j)^T B(e_i + e_j)$. This implies that $\alpha(n) \uparrow \alpha$. We conclude that $S_n(i, j)$ must converge to some $S(i, j)$. Hence $S_n(i, j) \to S(i, j)$.

(d) Simple algebraic manipulations show that

$$S = (A - ARC)^TS(A - ARC)^T + ARK_WRTA^T + K_V.$$

Assume that $(A - ARC)^Tv = \lambda v$ with $|\lambda| > 1$ and nonzero $v$. Then

$$v^*Sv = v^*(A - ARC)^TS(A - ARC)^Tv + v^*ARL_WRTA^Tv + v^*KK_Vv = |\lambda|^2v^*Sv + v^*ARL_WRTA^Tv + v^*KK_Vv.$$

Since $|\lambda| > 1$, it is obtained that $v^*Sv = v^*ARL_WRTA^Tv = v^*KK_Vv = v^*QQ^Tv = 0$, which implies $Qv = 0$ and $(A, Q)$ is not reachable.

(e) We borrow this proof from [3]. The notation is the same as before. Consider the system $\xi_{n+1} = A^T\xi_n + CT\xi_n + W_n$. The problem is to find the $z_k(\xi_k)$ for $k = 1, \ldots, n$ to achieve

$$\Phi_n(\xi) = \min_{\xi}E[\sum_{k=1}^{n-1} \xi_k^TK_W\xi_k + \xi_k^TK_W\xi_k + \xi_k^TS_1\xi_k | \xi_1 = \xi].$$

By considering all possible choices for $\xi_1$ and assuming that the cost is minimized over $\{2, \ldots, n + 1\}$, one sees that

$$\Phi_n(\xi) = \min_{\xi} \{\xi^TK_W\xi + \xi^TK_W\xi + E(\Phi_n(A^T\xi + CT\xi + W_1))\}.$$

A direct calculation (see below) allows to verify that

$$\Phi_n(\xi) = \xi^TS_n\xi + \sum_{k=1}^{n-1} \text{trace}(S_kK_V)$$

and the minimizing values of $\xi_n$ are given by

$$\xi_n = -[CS_kCT + K_W]^{-1}CS_kA^T\xi_{n-k}.$$

Instead of using this optimal control, one uses

$$\xi_n = -[CSCT + K_W]^{-1}CSAT\xi_{n-k}.$$
In the calculations above, we used the fact that, if $A_1$ is symmetric and invertible, then

$$\min_{\zeta} \{ \xi^T A_0 \xi + \zeta^T A_1 \zeta + 2 \xi^T A_2 \zeta \} = \xi^T [A_0 - A_2^T A_1^{-1} A_2] \xi$$

and is achieved by

$$\zeta = -A_1^{-1} A_2 \xi,$$

as a direct calculation shows.

(f) Assume that $S'$ is another positive semidefinite solution of (6). If $S_1 = S'$, then $S_n = S'$. However, we know from (e) that $S_n \to S$. Hence, $S' = S$.

(g) This follows immediately because the time-invariant filter is a special case of the time-varying filter.

REFERENCES