Proof of Lemma

Proof. The proof for this lemma follows the proof of theorem 6 in [1]. Here we modify the proof to account for the changing $C_t$ matrix. The proof is by induction. First, we prove that it holds for $t = 1$. Then we assume that it holds for $t = k$, we prove it holds for $t = k + 1$.

When $t = 1$, then

\[
\begin{align*}
\hat{x}_1 &= A\hat{x}_0 + K_0(y_t - C_0\hat{x}_0) = K_0y_t \\
K_0 &= (G_0 - AP_0C_0^T)(A_0^T - C_0P_0C_0^T)^{-1} = G_0(\Lambda_0)^{-1} = \Delta_0^0(L_1^0)^{-1} \\
\Rightarrow \hat{x}_1 &= \Delta_1^0(L_1^0)^{-1}y_t
\end{align*}
\]

When $t = 1$, we can also calculate $P_1$:

\[
P_1 = AP_0A^T + (G_0 - AP_0C_0^T)(A_0^T - C_0P_0C_0^T)^{-1}(G_0 - AP_0C_0^T)^T
\]

Hence, (13) and (14) hold for $t = 1$. Next we prove that if (13) and (14) hold for $t = k$, then they hold for $t = k + 1$. In this process, we will use the following formula for matrix inversion [2].

\[
\begin{pmatrix}
A_{11} & A_{12} \\
A_{21} & A_{22}
\end{pmatrix}^{-1} = 
\begin{pmatrix}
A_{11}^{-1} + A_{11}^{-1}A_{12}(A_{22} - A_{21}A_{11}^{-1}A_{12})^{-1}A_{21}A_{11}^{-1} & -A_{11}^{-1}A_{12}(A_{22} - A_{21}A_{11}^{-1}A_{12})^{-1} \\
-A_{22} + A_{21}A_{11}^{-1}A_{12}^{-1}A_{21}A_{11}^{-1} & (A_{22} - A_{21}A_{11}^{-1}A_{12})^{-1}
\end{pmatrix}
\]

Let’s calculate $P_{k+1}$ first:

\[
P_{k+1} = \Delta_{k+1}^0(L_{k+1}^0)^{-1}(\Delta_{k+1}^0)^T
\]

\[
= (A\Delta_k^0 G_k) \left( L_k^0 \begin{pmatrix}
A_k^0 & 0 \\
A_k^0 & A_k^0
\end{pmatrix} \right)^{-1} \left( \begin{pmatrix}
\Delta_k^0 & 0 \\
\Delta_k^0 & \Delta_k^0
\end{pmatrix} \right)
\]

\[
= (A\Delta_k^0 G_k) \left( (L_k^0)^{-1} + (L_k^0)^{-1}(\Delta_k^0)^T C_k^T \Delta_k^{-1} \Delta_k \Delta_k^0 (L_k^0)^{-1} \right) \left( \begin{pmatrix}
\Delta_k^0 & 0 \\
\Delta_k^0 & \Delta_k^0
\end{pmatrix} \right) \left( \begin{pmatrix}
\Delta_k^0 & 0 \\
\Delta_k^0 & \Delta_k^0
\end{pmatrix} \right)
\]

where $\Delta = \Lambda_k^0 - C_k\Delta_k^0 (L_k^0)^{-1}(\Delta_k^0)^T C_k^T = \Lambda_k^0 - C_kP_kC_k^T$

\[
= \Delta_k^0(L_k^0)^{-1}(\Delta_k^0)^T A_k^T + (G_k - \Delta_k^0(L_k^0)^{-1}(\Delta_k^0)^T C_k^T)\Delta^{-1}(G_k - \Delta_k^0(L_k^0)^{-1}(\Delta_k^0)^T C_k^T)^T
\]

\[
= AP_kA^T + (G_k - AP_kC_k^T)(\Lambda_k^0 - C_kP_kC_k^T)^{-1}(G_k - AP_kC_k^T)^T
\]
This proves (14). Next we will prove (13).

\[
\hat{x}_{k+1} = \Delta^0_{k+1}(L^0_{k+1})^{-1} \begin{pmatrix} y_0 \\ y_1 \\ \vdots \\ y_k \end{pmatrix}
\]

\[
= (A\Delta^0_k \ G_k) \left( L^0_k \ A^T_k \right) \begin{pmatrix} y_0 \\ y_1 \\ \vdots \\ y_k \end{pmatrix}
\]

\[
= (A\Delta^0_k \ G_k)
\]

\[
\left( (L^0_k)^{-1} + (L^0_k)^{-1}(\Delta^0_k)^T C_k^T \Delta^{-1} C_k \Delta^0_k(L^0_k)^{-1} \right) \begin{pmatrix} y_0 \\ y_1 \\ \vdots \\ y_k \end{pmatrix}
\]

\[
= [A - (G_k - A\Delta^0_k(L^0_k)^{-1}(\Delta^0_k)^T C_k^T \Delta^{-1} C_k \Delta^0_k(L^0_k)^{-1})]
\]

\[
\hat{x}_{k} + (G_k - A\Delta^0_k(L^0_k)^{-1}(\Delta^0_k)^T C_k^T \Delta^{-1} C_k \Delta^0_k(L^0_k)^{-1})(y_k - C_k\hat{x}_k)
\]

\[
= A\hat{x}_k + K_k(y_k - C_k\hat{x}_k)
\]

This proves equation (13). \(\square\)

**Proof of Theorem 1**

**Proof.** From the definition of \(H^d_t\) and the ergodicity of \(y_t\), as \(N \to \infty\),

\[
H^d_t = E[Y_{t+d,r}Y_t^T]E[Y_{t+d-1}Y_{t,d-1}] = N^d_t(L^d_t)^{-1}Y_{t,d-1}
\]

It is easy to verify that

\[
N^d_t = O_{t,r|d}\Delta^d_t
\]

Hence, we have

\[
H^d_t = O_{t,r|d}\Delta^d_t(L^d_t)^{-1}Y_{t,d-1}
\]

By equation (15), we can write

\[
H^d_t = O_{t,r|d}\hat{X}_{t,r}
\]

Since \(O_{t,r|d}\) is full column rank, \(H^d_t\) has the same row space with \(\hat{X}_{t,r}\). The pair \((A, Q^{1/2})\) being controllable can guarantee that \(\hat{X}_{t,r}\) is full row rank, so \(H^d_t\) has the same column space with \(O_{t,r|d}\). So we can do singular value decomposition for \(H^d_t\) and find its row space and column space, respectively. Then we can get \(\hat{O}_{t,r|d}\) and \(\hat{X}_{t,r}^{\text{red}}\) up to a similarity transformation. \(\square\)

**Proof of Theorem 2**

**Proof.** By (21), we have

\[
\hat{O}_{t,1,1}^{\text{ref}}G_t = O_{t,1,1}^{\text{ref}}, \quad \hat{O}_{t+1,1,1}^{\text{ref}}G_{t+1} = O_{t+1,1,1}^{\text{ref}}
\]

So we have the following equation:

\[
\hat{O}_{t}^{\text{ref}} = O_t^{\text{ref}}, \quad \hat{O}_{t+1}^{\text{ref}}G_{t+1} = O_{t+1}^{\text{ref}} = O_t^{\text{ref}}
\]

By the above equation, we have

\[
\hat{O}_{t}^{\text{ref}}G_t = \hat{O}_{t+1}^{\text{ref}}G_{t+1} \Rightarrow G_{t+1}G_t^{-1} = (\hat{O}_{t+1}^{\text{ref}})^{\dagger}\hat{O}_{t}^{\text{ref}}
\]
That is, the transformation matrix between $\hat{O}_{t+1}^{\text{ref}}$ and $\hat{O}_t^{\text{ref}}$ is given by $\Gamma = \Gamma_{t+1} \Gamma_t^{-1}$. Hence, $\hat{O}_{t+1,1|d} \Gamma$ and $\hat{O}_{t,1|d}$ will stay in the same basis. So we have

$$\hat{A} = (\hat{O}_{t+1,1|d} \Gamma)^\dagger \hat{O}_{t,1|d}$$

Proof of Theorem 3

Proof. The proof is straightforward. To prove this theorem, we use a concrete example. Let’s consider the case of using three sensors to collect measurements in Figure 1. In this case, $N_s = 2$, $N_c = 3$. Let $r = 3$. So all the three sensors will execute the motion that they move forward for one step and stay for $r = 3$ steps and move again. Without loss of generality, let’s assume that they start at time step $t = 1$ and sensor 1 starts at waypoint 1, sensor 2 starts at waypoint 2, and sensor 3 starts at waypoint 3. When $K_d = 2$, the corresponding extended observability matrix is given by

$$O_{1,3|6}^{\text{mult}} = \begin{pmatrix} C_1^{\text{mult}} \\ C_2^{\text{mult}} A \\ C_3^{\text{mult}} A^2 \\ C_4^{\text{mult}} A^3 \\ C_5^{\text{mult}} A^4 \\ C_6^{\text{mult}} A^5 \end{pmatrix}, \quad \hat{O}_{1,3|6}^{\text{mult}} = \begin{pmatrix} C_1^{\text{mult}} \\ C_2^{\text{mult}} A \\ C_3^{\text{mult}} A^2 \\ C_4^{\text{mult}} A^3 \\ C_5^{\text{mult}} A^4 \\ C_6^{\text{mult}} A^5 \end{pmatrix}, \quad Q_{1,3|6}^{\text{mult}} = \begin{pmatrix} C_2^{\text{mult}} A \\ C_3^{\text{mult}} A^2 \\ C_4^{\text{mult}} A^3 \\ C_5^{\text{mult}} A^4 \\ C_6^{\text{mult}} A^5 \end{pmatrix}$$

(1)

where $C_t^{\text{mult}} = [C_1^t, C_2^t, C_3^t]$ for all $t = 1, 2, \ldots, 6$. And $C_1^{\text{mult}} = C_2^{\text{mult}} = C_3^{\text{mult}} = C_4^{\text{mult}} = C_5^{\text{mult}} = C_6^{\text{mult}}$. We also have $C_2^t = C_1^t$ and $C_3^t = C_2^t$ since sensor 1 and sensor 2 are consecutive pairs 1 $\rightarrow$ 2 and sensor 2 and sensor 3 are also consecutive pairs 2 $\rightarrow$ 3. Hence, the number of rows in $O_{1,3|6}^{\text{mult}}$ is 14, which is equal to $N_s (r-1) K_d + N_c (K_d - 1)$ when $N_s = 3, r = 3, K_d = 2, N_c = 2$. The $N_s (r-1) K_d$ rows come from $[C_1^{\text{mult}}, C_2^{\text{mult}} A; C_4^{\text{mult}} A^3; C_5^{\text{mult}} A^4]$ and the $N_c (K_d - 1)$ rows come from $[C_2^t A^2; C_3^t A^2]$. This proof can be generalized to any $N_s$, $N_c$, $r$ and $K_d$ except that $N_s = 1$ and $r = 1$ at the same time. In this case, we only have one sensor and it keeps moving along the periodic trajectory and $N_c = 0$. We need to use a static reference sensor to learn the $A$ matrix and $K_d \geq n + 1$.}

References