Date. 06., June. 2021.

DISCUSSION OF ANALYTICAL SOLUTION OF LINEAR MHE.

Reference: H. Cox "Estimation of State variables for noisy dynamic systems" Ph. D Thesis. M. I. T.

Summary:

For linear systems, there is no assumptions or approximations on the derivation of the analytical solution.

The following terms (states) are defined to ease the derivation and make the expression compact.

$$\hat{x}(k|k-1) \triangleq F(k)\cdot\hat{x}(k-1|k-1)$$

$$\hat{x}(k|k) \triangleq \hat{x}(k|k-1) + C(k)\cdot H^T(k) R^{-1}(k)\left[Z(k) - H(k)\cdot\hat{x}(k|k-1)\right]$$

The right-hand side of the above definition is generated in the derivation. Therefore, the solution with the above definition is identical to that obtained directly from static optimization. We refer to $\hat{x}(k|k-1)$ as one step prediction based on the last estimate, and $\hat{x}(k|k)$ as the estimate for current time.

We justify the definition by considering two special cases and then ~~generalizing~~ ~~to~~ extending it to general case. This is to show that the above definition is derived using PMP or KKT necessary optimality conditions.

~~Case I n=0.~~

Let us consider the general MHE problem of the following form:

$$\min_{\{x_k\},\{w(k)\}} J = \frac{1}{2}\|x(0) - m\|^2_{P^{-1}} + \sum_{k=0}^{n} \frac{1}{2}\|Z(k) - H(k)\cdot x(k)\|^2_{R^{-1}(k)} + \sum_{k=0}^{n-1}\frac{1}{2}\|w(k)\|^2_{Q^{-1}(k)}$$

$$\text{s.t.} \quad x(k+1) = F(k)\cdot x(k) + G(k)\cdot w(k) \qquad k=0,\cdots,n-1.$$

The decision variables are $x(k|n)$, and $w(k)$. $x(k|n)$ is the smooth at time k based on the observation sequence $\{Z(0),\cdots,Z(n)\}$.

*. Eq. (1) is similar to Kalman Filter.

$$\hat{x}(0|0) = m + \underbrace{P(0)\cdot H^T(0)\cdot R^{-1}(0)}_{\text{gain}}\underbrace{[z(0) - H(0)\cdot\hat{x}(0|0)]}_{\text{correction}}$$

Date. 06 June. 2021

In this sense, the optimization problem of MHE may be re-formulated as:

$$\min_{\{x(k|n)\}\{w(k)\}} J = \frac{1}{2}\|\hat{x}(0|n) - m\|^2_{P^{-1}} + \sum_{k=0}^{n}\frac{1}{2}\|z(k) - H(k)\cdot\hat{x}(k|n)\|^2_{R^{-1}(k)} + \frac{1}{2}\sum_{k=0}^{n-1}\|w(k)\|^2_{Q^{-1}(k)}$$

$$\text{s.t.} \quad \hat{x}(k+1|n) = F(k)\cdot\hat{x}(k|n) + G(k)\cdot w(k) \quad k=0,\cdots,n-1.$$

Case I. $n=0$, in which case the decision variable is $\hat{x}(0|0)$ only.

In this case, the PMP conditions give the following equations.

$$\hat{x}(0|0) = m + P(0)H^T(0)R^{-1}(0)[z(0) - H(0)\cdot\hat{x}(0|0)] \quad \text{-----} \quad (1).$$

This is achieved by setting the gradient of the Lagrangian equal to zero.

The corresponding Lagrangian is:

$$L = \frac{1}{2}\|\hat{x}(0|0) - m\|^2_{P^{-1}} + \frac{1}{2}\|z(0) - H(0)\cdot\hat{x}(0|0)\|^2_{R^{-1}}$$

Note that there is no constraint so that $\lambda(0) = 0$, which agrees with $\lambda(n) = 0$.

In this special case, $\hat{x}(0|0)$ can be interpreted as the estimate at current time.

Case II. $n=1$, in which case the decision variables are $\{\hat{x}(0|1), \hat{x}(1|1), w(0)\}$.

The corresponding Lagrangian is:

$$L = \frac{1}{2}\|\hat{x}(0|1) - m\|^2_{P^{-1}} + \frac{1}{2}\|z(0) - H(0)\cdot\hat{x}(0|1)\|^2_{R^{-1}(0)} + \frac{1}{2}\|z(1) - H(1)\cdot\hat{x}(1|1)\|^2_{R^{-1}(1)} + \frac{1}{2}\|w(0)\|^2_{Q^{-1}(0)} + \lambda^T(0)\cdot[\hat{x}(1|1) - F(0)\cdot\hat{x}(0|1) - G(0)\cdot w(0)]$$

KKT conditions:

Primal feasibility: $\frac{\partial L}{\partial\lambda(0)} = \hat{x}(1|1) - F(0)\hat{x}(0|1) - G(0)\cdot w(0) = 0$

Zero-gradient of Lagrangian: $-P^{-1}(0)$ → Boundary condition

$$x \triangleq \begin{bmatrix}\hat{x}(0|1)\\ \hat{x}(1|1)\\ w(0)\end{bmatrix} \quad \frac{\partial L}{\partial x} = \begin{bmatrix}\frac{\partial L}{\partial\hat{x}(0|1)}\\ \frac{\partial L}{\partial\hat{x}(1|1)}\\ \frac{\partial L}{\partial w(0)}\end{bmatrix} = \begin{bmatrix}m + P(0)\cdot H^T(0)\cdot R^{-1}(0)[z(0) - H(0)\cdot\hat{x}(0|1)] + P(0)\cdot F^T(0)\cdot\lambda(0)] + P^{-1}(0)\cdot\hat{x}(0|1)\\ \lambda(0) - H^T(1)\cdot R^{-1}(1)[z(1) - H(1)\cdot\hat{x}(1|1)]\\ Q^{-1}(0)\cdot w(0) - G^T(0)\cdot\lambda(0)\end{bmatrix} = \begin{bmatrix}0\\ 0\\ 0\end{bmatrix} \begin{matrix}(2)\\ (3)\\ (4)\end{matrix}$$

Large optimization variable vector.

☆ Dynamic programming is converted to a large static optimization problem.

Date. 06 June. 2021

In what follows, we will express the KKT conditions of $n=1$ in terms of $\hat{x}(0|0)$, $\hat{x}(0|0)$ that were defined in Page 312.

Step 1. Rewrite the optimal estimate (1) as below:

$$\hat{x}(0|0) = m + P(0) H^T(0) R^{-1}(0) [Z(0) - H(0)\cdot\hat{x}(0|0)]$$
$$= m + P(0)\cdot H^T(0)\cdot R^{-1}(0)\cdot H(0)\cdot m - P(0)\cdot H^T(0) R^{-1}(0) H(0)\cdot m$$
$$+ P(0)\cdot H^T(0)\cdot R^{-1}(0) Z(0) - P(0)\cdot H^T(0)\cdot R^{-1}(0)\cdot H(0)\cdot\hat{x}(0|0).$$
$$\Downarrow$$
$$[I + P(0)\cdot H^T(0) R^{-1}(0) H(0)]\,\hat{x}(0|0) = [I + P(0) H^T(0) R^{-1}(0) H(0)] m + P(0)\cdot H^T(0)\cdot R^{-1}(0)\cdot [Z(0) - H(0)\cdot m]$$
$$\Downarrow$$
$$\hat{x}(0|0) = m + \underbrace{[I + P(0)\cdot H^T(0)\cdot R^{-1}(0)\cdot H(0)]^{-1}\cdot P(0)}_{C(0)\ \text{(defined as } C(0))}\cdot H^T(0)\cdot R^{-1}(0)\cdot [Z(0) - H(0)\cdot m]$$
$$\Downarrow$$
$$\hat{x}(0|0) = m + C(0)\cdot H^T(0)\cdot R^{-1}(0)\cdot [Z(0) - H(0)\cdot m] \quad \cdots\cdots (5)$$

Step 2. Rewrite the boundary condition (2) using the same trick as above.

$$\hat{x}(0|1) = m + P(0)\cdot H^T(0)\cdot R^{-1}(0) [Z(0) - H(0)\cdot\hat{x}(0|1)] + P(0)\cdot F^T(0)\cdot\lambda(0)$$
$$\hat{x}(0|1) = m + P(0)\cdot H^T(0) R^{-1}(0)\cdot H(0) m - P(0)\cdot H^T(0) R^{-1}(0)\cdot H(0) m$$
$$+ P(0)\cdot H^T(0)\cdot R^{-1}(0) Z(0) - P(0)\cdot H^T(0)\cdot R^{-1}(0)\cdot H(0)\cdot\hat{x}(0|1) + P(0)\cdot F^T(0)\cdot\lambda(0)$$
$$[I + P(0)\cdot H^T(0) R^{-1}(0) H(0)]\,\hat{x}(0|1) = [I + P(0)\cdot H^T(0) R^{-1}(0)\cdot H(0)] m + P(0) H^T(0) R^{-1}(0) [Z(0) - H(0)\cdot m]$$
$$+ P(0)\cdot F^T(0)\cdot\lambda(0)$$
$$\Downarrow$$
$$\hat{x}(0|1) = m + [I + P(0)\cdot H^T(0)\cdot R^{-1}(0)\cdot H(0)]^{-1} P(0)\cdot H^T(0)\cdot R^{-1}(0) [Z(0) - H(0)\cdot m]$$
$$+ [I + P(0)\cdot H^T(0) R^{-1}(0)\cdot H(0)]^{-1}\cdot P(0)\cdot F^T(0)\cdot\lambda(0)$$
$$\Downarrow \quad \text{With the definition of } C(0)$$

* 5-Sept-2021 This is exactly obtained by setting $n=0$.

$$\hat{x}(0|1) = m + C(0)\cdot H^T(0)\cdot R^{-1}(0) [Z(0) - H(0)\cdot m] + C(0)\cdot F^T(0)\cdot\lambda(0) \quad \cdots (6)$$

Comparison with (5):

$$\hat{x}(0|0) = m + C(0)\cdot H^T(0)\cdot R^{-1}(0) [Z(0) - H(0)\cdot m]$$

gives: $\hat{x}(0|1) = \hat{x}(0|0) + C(0)\cdot F^T(0)\cdot\lambda(0) \quad \cdots\cdots (7)$

This constructs a relationship between $\hat{x}(k|n)$ and $\hat{x}(k|k)$.

So, in general, the smooth, i.e. the decision variables, can be obtained by the estimate according to the ~~ex~~recursive relationship:

$$\hat{x}(k|n) = \hat{x}(k|k) + C(k)\cdot F^T(k)\cdot\lambda(k)$$

See the proof in page D1 of "Draft 2".

Date. 06. June. 2021

Another important observation is that the equation (5) can be generalized in any time instant k in the following form. why?

$$\hat{x}(k|k) = \underbrace{\hat{x}(k|k-1)}_{\text{One step prediction}} + \underbrace{\boxed{C(k) H^T(k) R^{-1}(k)}}_{\text{Kalman Gain}^*} \underbrace{[Z(k) - H(k)\cdot \hat{x}(k|k-1)]}_{\text{Correction}}$$

This is exactly the Kalman Filter! The optimal estimation of current state.

Step 3. Rewrite the primal feasibility in the form of Kalman Filter using the zero-gradient of Lagrangian and $\boxed{\text{the equation (7)}}$

Primal feasibility: $\hat{x}(1|1) = F(0)\cdot \hat{x}(0|1) + G(0)\cdot W(0)$

Plugging $W(0) = Q(0)\cdot G^T(0)\cdot \lambda(0)$ obtained from (4) to the above equation yields: $\hat{x}(1|1) = F(0)\cdot \hat{x}(0|1) + G(0)\cdot Q(0)\cdot G^T(0)\cdot \lambda(0)$ ------- (8)

With the relationship of $\boxed{\hat{x}(0|1) = \hat{x}(0|0) + C(0)\cdot F^T(0)\,\lambda(0)}$ we can rewrite eq (8) as: $\hat{x}(1|1) = F(0)\cdot \hat{x}(0|0) + \underbrace{[F(0)\cdot C(0)\cdot F^T(0) + G(0)\cdot Q(0)\cdot G^T(0)]}_{\text{defined as } P(1)}\lambda(0)$ ... (9)

This is the key that relates $\hat{x}(k|k)$ with $\hat{x}(k|k-1)$

Plugging $\lambda(0) = H^T(1) R^{-1}(1)[Z(1) - H(1)\cdot \hat{x}(1|1)]$ obtained from (3) to eq. (9) yields: $\hat{x}(1|1) = \underbrace{F(0)\cdot \hat{x}(0|0)}_{\text{defined as } \hat{x}(1|0)} + P(1)\cdot H^T(1)\cdot R^{-1}(1)\cdot [Z(1) - H(1)\cdot \hat{x}(1|1)]$

← This definition does not assume any approximations.

So far, we have not made any approximations and assumptions, all the derivations are directly based on algebraic skills and techniques.

$$\begin{aligned}\hat{x}(1|1) &= \hat{x}(1|0) + P(1)\cdot H^T(1)\cdot R^{-1}(1)[Z(1) - H(1)\cdot \hat{x}(1|1)] \\ &= \hat{x}(1|0) + P(1)\cdot H^T(1)\cdot R^{-1}(1)\cdot H(1)\cdot \hat{x}(1|0) - P(1) H^T(1) R^{-1}(1) H(1)\cdot \hat{x}(1|0) \\ &\quad + P(1) H^T(1) R^{-1}(1) Z(1) - P(1)\cdot H^T(1) R^{-1}(1)\cdot H(1)\hat{x}(1|1)\end{aligned}$$

⇓

$$[I + P(1)\cdot H^T(1)\cdot R^{-1}(1)\cdot H(1)]\,\hat{x}(1|1) = [I + P(1)\cdot H^T(1)\cdot R^{-1}(1)\cdot H(1)]\,\hat{x}(1|0) + P(1) H^T(1) R^{-1}(1)[Z(1) - H(1)\cdot \hat{x}(1|0)]$$

$$\hat{x}(1|1) = \hat{x}(1|0) + \underbrace{[I + P(1) H^T(1) R^{-1}(1) H(1)]^{-1} P(1)}_{\text{defined as } C(1)} H^T(1) R^{-1}(1)[Z(1) - H(1)\cdot \hat{x}(1|0)]$$

⇓

$$\hat{x}(1|1) = \hat{x}(1|0) + C(1)\cdot H^T(1)\cdot R^{-1}(1)[Z(1) - H(1)\cdot \hat{x}(1|0)]$$

which is exactly the form of Kalman Filter.

* 13. June. 2021. The Kalman Gain is obtained from $C(k)$ using the matrix inversion lemma.

Date. 06. June. 2021.

The analytical solution of $n=1$ is summarized below.

$$\hat{x}(0|1) = \hat{x}(0|0) + C(0)\cdot F^T(0)\cdot \lambda(0).$$

$\cancel{w(0)} \quad$ where $\hat{x}(0|0)$ is given

Kalman Filter $\Rightarrow$ $$\hat{x}(1|1) = \hat{x}(1|0) + C(1)\cdot H^T(1)\cdot R^{-1}(1)\cdot[z(1) - H(1)\cdot\hat{x}(1|0)]$$

where $\hat{x}(1|0) = F(0)\cdot \hat{x}(0|0)$

$$\lambda(0) = H^T(1)\cdot R^{-1}(1)[z(1) - H(1)\hat{x}(1|1)]$$

$$w(0) = Q(0)\cdot G^T(0)\cdot \lambda(0)$$

Compute (Solve) the Kalman Filter form, then compute the costate, followed by computing $\hat{x}(0|1)$, finally, compute $w(0)$.

The obtained closed-form solution is easy to implement in computer as it has the recursive form. Otherwise, we have to resort to the following linear matrix which is difficult to solve, especially for high dimension & (large horizon)

$$\underbrace{\begin{bmatrix} I, & -F(0), & -G(0), & 0 \\ P^{-1}(0)+H^T(0)R^{-1}(0)H(0), & 0, & 0, & -F^T(0) \\ 0, & H^T(1)R^{-1}(1)H(1), & 0, & I \\ 0, & 0, & Q^{-1}(0), & -G^T(0) \end{bmatrix}}_{\text{KKT matrix}} \begin{bmatrix} \hat{x}(0|1) \\ \hat{x}(1|1) \\ w(0) \\ \lambda(0) \end{bmatrix} = \begin{bmatrix} 0 \\ P^{-1}(0)m + H^T(0)R^{-1}(0)H(0) \\ H^T(1)R^{-1}(1)z(1) \\ 0 \end{bmatrix}$$

Next, we show that the recursive relationship $\hat{x}(k|n) = \hat{x}(k|k) + C(k)F^T(k)\lambda(k)$ holds for arbitrary k, of course, it also applies to MHE.

proof. The proof is by induction. We are going to show the relationship holds for $k+1$, namely, $\hat{x}(k+1|n) = \hat{x}(k+1|k+1) + C(k+1)F^T(k+1)\lambda(k+1)$

1. Primal feasibility at $k+1$:

$$\hat{x}(k+1|n) = F(k)\cdot\hat{x}(k|n) + G(k)\cdot w(k)$$

2. Zero-gradient w.r.t w_k:

$$w_k = Q(k)\cdot G^T(k)\cdot \lambda(k)$$

Combining 1 and 2, we have:

* 22. June. 2021.
The conversion to a standard Kalman filter is presented in 'Draft 2', p2. (D1)

thanks to this linear relationship we can derive analytic solution.

Date. 06. June. 2021

$$\hat{x}(k+1|n) = F(k)\cdot\hat{x}(k|n) + G(k)\cdot Q(k)\cdot G^T(k)\cdot\lambda(k) \quad \text{---- (1)}$$

3. Recursive relationship at k

$$\hat{x}(k|n) = \hat{x}(k|k) + C(k)\cdot F^T(k)\cdot\lambda(k)$$

Plugging the above relationship to (1) yields:

$$\hat{x}(k+1|n) = \underbrace{F(k)\cdot\hat{x}(k|k)}_{\text{defined as } \hat{x}(k+1|k)} + \underbrace{\left[F(k)\cdot C(k)\cdot F^T(k) + G(k)Q(k)\cdot G^T(k)\right]}_{\text{defined as } P(k+1)}\lambda(k) \quad \text{--- (2)}$$

corresponding to covariance of prediction error in KF
13. Sept. 2021.

$$\Downarrow$$

$$\hat{x}(k+1|n) = \hat{x}(k+1|k) + P(k+1)\lambda(k) \quad \text{------ (3)}$$

4. Zero-gradient w.r.t $\hat{x}(k+1|n)$:

$$\lambda(k) = F^T(k+1)\lambda(k+1) + H^T(k+1)R^{-1}(k+1)\left[z(k+1) - H(k+1)\hat{x}(k+1|n)\right]$$

Plugging the expression of $\lambda(k)$ in (3) yields:

$$\hat{x}(k+1|n) = \hat{x}(k+1|k) + P(k+1)H^T(k+1)R^{-1}(k+1)\left[z(k+1) - H(k+1)\hat{x}(k+1|n)\right] + P(k+1)F^T(k+1)\lambda(k+1)$$

$$\Downarrow$$

$$\begin{aligned}\hat{x}(k+1|n) = {} & \hat{x}(k+1|k) + P(k+1)H^T(k+1)R^{-1}(k+1)H(k+1)\hat{x}(k+1|k) \\ & - P(k+1)H^T(k+1)R^{-1}(k+1)H(k+1)\hat{x}(k+1|k) \\ & + P(k+1)H^T(k+1)R^{-1}(k+1)z(k+1) - P(k+1)H^T(k+1)R^{-1}(k+1)H(k+1)\hat{x}(k+1|n) \\ & + P(k+1)F^T(k+1)\lambda(k+1)\end{aligned}$$

$$\Downarrow$$

corresponding to covariance of estimation error in KF.

$$\begin{aligned}\hat{x}(k+1|n) = {} & \hat{x}(k+1|k) + \underbrace{\left[I + P(k+1)H^T(k+1)R^{-1}(k+1)H(k+1)\right]^{-1}P(k+1)}_{\text{defined as } C(k+1)}H^T(k+1)R^{-1}(k+1) \\ & \left[z(k+1) - H(k+1)\hat{x}(k+1|k)\right] \\ & + \left[I + P(k+1)H^T(k+1)R^{-1}(k+1)H(k+1)\right]^{-1}P(k+1)F^T(k+1)\lambda(k+1)\end{aligned}$$

$$\Downarrow$$

$$\hat{x}(k+1|n) = \underbrace{\hat{x}(k+1|k) + C(k+1)\cancel{P(k+1)}H^T(k+1)R^{-1}(k+1)\left[z(k+1) - H(k+1)\hat{x}(k+1|k)\right]}_{\hat{x}(k+1|k+1) \text{ due to the Kalman Filter.}} + C(k+1)\cdot F^T(k+1)\cdot\lambda(k+1)$$

Finally, we obtain the desired relationship for $k+1$.

$$\hat{x}(k+1|n) = \hat{x}(k+1|k+1) + C(k+1)F^T(k+1)\cdot\lambda(k+1)$$

All the steps in this proof are directly based on KKT conditions.

This completes the proof.