Learning 3DVar

Lecture 4

GOALS //

- Define Online State Estimation
- Define 3DVAR
- Learning The Gain
- Connection to Kalman Filter

ONLINE STATE ESTIMATION //

ⓐ $v^{+}_{j+1} = \Psi(v^{+}_{j}) + \xi^{+}_{j}$, $j \in \mathbb{Z}^{+}$; $v^{+}_{0} \sim N(m_0, C_0)$

ⓑ $y^{+}_{j+1} = Hv^{+}_{j+1} + \eta^{+}_{j+1}$, $j \in \mathbb{N}$

$\xi^{+}_{j} \sim N(0,\Sigma)$, $\eta^{+}_{j+1} \sim N(0,\Gamma)$ i.i.d $\quad \eta_j \perp\!\!\!\perp v^{+}_{0} \perp\!\!\!\perp \xi_j$

Assumption / $\{y^{+}_{1}, y^{+}_{2}, \ldots\}$ arrive sequentially

Estimate / v^{+}_{j} sequentially.

Notation / $\quad v_j$ estimates v^{+}_{j}

$$Y^{+}_{j} = \{y^{+}_{1}, \ldots, y^{+}_{j}\}$$

$$v_j = \mathcal{F}_0(v_0, Y^{+}_{j+1})$$

Markov / $\quad v_{j+1} = \mathcal{F}_1(v_j, y^{+}_{j+1})$

3DVAR / $\hat{v}_{j+1} = \Psi(v_j)$ Predict

$$v_{j+1} = \underset{v}{\operatorname{argmin}}\ J_j(v) \quad \text{optimize}$$

$$J_j(v) = \frac{1}{2}\|\hat{v}_{j+1} - v\|^2_{\hat{C}} + \frac{1}{2}\|y^{+}_{j+1} - Hv\|^2_{\Gamma}$$

Lemma / $v_{j+1} = (I - KH)\Psi(v_j) + K y^{+}_{j+1}$,

$$K = (\hat{C}^{-1} + H^T\Gamma^{-1}H)^{-1} H^T \Gamma^{-1}.$$

Proof /

$$(\hat{C}^{-1} + H^T\Gamma^{-1}H)\, v_{j+1} = \hat{C}^{-1}\hat{v}_{j+1} + H^T\Gamma^{-1}y^{+}_{j+1}$$

Define $C = (\hat{C}^{-1} + H^T\Gamma^{-1}H)^{-1}$. Then

Invert in state space

$$C^{-1}v_{j+1} = C^{-1}\hat{v}_{j+1} - H^T\Gamma^{-1}H\hat{v}_{j+1} + H^T\Gamma^{-1}y^{+}_{j+1}$$

$$v_{j+1} = \hat{v}_{j+1} + CH^T\Gamma^{-1}(y^{+}_{j+1} - H\hat{v}_{j+1})$$

Result follows since $\hat{v}_{j+1} = \Psi(v_j)$.

Remark / We know Γ but not $\hat{C}$. Learn $\hat{C}$?

Better to learn K. (the Gain)

LEARNING THE GAIN // Recall algorithm:

$$v_{j+1} = (I - KH)\,\Psi(v_j) + K y^\dagger_{j+1}, \quad v_0 = m_0$$

$$\boxed{\text{Data Assumption / we are given } \{v_j^\dagger, y_j^\dagger\}_{j \in \{1,\dots,J\}}}$$

$*$

Note that $v_j = v_j(K)$. Thus define

$$J(K) = \frac{1}{J}\sum_{j=1}^{J} \| v_j(K) - v_j^\dagger \|^2$$

Note that $v_j(K) = v_j(K, Y_j^\dagger)$.

we can evaluate $J(\cdot)$ under Data Assumption. $K^* = \operatorname{argmin} J^J(K)$

Remark/ using Sherman-Woodbury formula:

$$K = \hat{C}H^T(H\hat{C}H^T + \Gamma)^{-1}$$

Invert in data space.

CONNECTION TO KALMAN FILTER // $\Psi(\cdot) = A(\cdot)$

$$\left.\begin{aligned} v^\dagger_{j+1} &= A v_j^\dagger + \xi_j^\dagger \\ y^\dagger_{j+1} &= H v^\dagger_{j+1} + \eta^\dagger_{j+1} \end{aligned}\right\} \text{ Then } \mathbb{P}(v_j^\dagger \mid Y_j^\dagger) = N(v_j, C_j)$$

$$\hat{v}_{j+1} = A v_j$$

$$v_{j+1} = (I - K^{kal}_{j+1} H)\hat{v}_{j+1} + K^{kal}_{j+1} y_{j+1}$$

Mean update

$$\hat{C}^{kal}_{j+1} = A C^{kal}_j A^T + \Sigma$$

$$K^{kal}_{j+1} = \hat{C}^{kal}_{j+1} H^T (H \hat{C}^{kal}_{j+1} H^T + \Gamma)^{-1}$$

$$C^{kal}_{j+1} = (I - K^{kal}_{j+1} H)\hat{C}^{kal}_{j+1}$$

cov & Kalman update

Thus mean of Kalman filter evolves similarly to 3DVAR but with adaptive gain.

However:

Theorem/ Assume that

$$(C^{kal}_j, \hat{C}^{kal}_j, K^{kal}_j) \to (C_\infty, \hat{C}_\infty, K_\infty)$$

as $j \to \infty$. Then define

$$J_\infty(K) = \lim_{J\to\infty} J(K).$$

It follows that $K_\infty = \underset{K}{\operatorname{argmin}}\ J_\infty(K)$

From Kalman Filter

Learned