Data Assimilation 1: Analytical Methods

Lesson 9
FORECAST-ANALYSIS CYCLE

Predict the future using your current understanding of the system
- Model-based
- Error-propagation

Scientific method cycle

Update

Update prior understanding of the system based on new information

Current State
Future state
Observations
The Analysis Problem

• Prior to observing how the future plays out, what is our best estimate of the future state of the system, \(X_{t+1}\)?

• The forecast, \(P(X_{t+1})\)

• Once we make (imperfect) observations of the system, \(Y_t\), what’s our best estimate of \(X_t\)?

• \(P(X_{t+1}) = P(Y_{t+1})\) ?

• \(P(X_{t+1} | Y_{t+1}) \propto P(Y_{t+1} | X_{t+1}) P(X_{t+1})\)

\(\text{Prior} \quad \text{Likelihood} \quad \text{Posterior}\)
Simplest Analysis

- Forecast:
  Assume $P(X_{t+1}) \sim N(\mu_f, p_f)$

- Observation error:
  Assume $P(Y_{t+1} \mid X_{t+1}) \sim N(X_{t+1}, r)$

- Likelihood = Data model

- Assume $Y, \mu_f, p_f$ and $r$ are known

- $P(X_{t+1} \mid Y_{t+1}) \sim N(\mu_a, p_a)$
  \[ \rho = 1/r \quad \phi = 1/p_f \]

- $X \mid Y \sim N \left( \frac{\rho}{n\rho + \phi} n\bar{Y} + \frac{\phi}{n\rho + \phi} \mu_f, n\rho + \phi \right)$
Precision controls influence

Less Precise Data

Less Precise Model
Simplest Forecast

- Process Model
  \[ X_{t+1} = mX_t + \varepsilon_t \]

- Process error
  \[ \varepsilon_t \sim N(0,q) \]

- Assume \( m \) and \( q \) are known

- State uncertainty (IC)
  \[ P(X_t \mid Y_t) \sim N(\mu_a,p_a) \]

- What is \( P(X_{t+1}) \)?

- \[ E[X_{t+1}] = E[mX_t + \varepsilon_t] = m\mu_a \]

- \[ \text{Var}[X_{t+1}] = \text{Var}[mX_t + \varepsilon_t] \]
  \[ \approx m^2\text{Var}[X_t] + \text{Var}[\varepsilon_t] - 2\text{Cov}[mX_t,\varepsilon_t] \]

- \[ \approx m^2\text{Var}[X_t] + \text{Var}[\varepsilon_t] \]

- \[ m^2p_a + q \]

- \[ P(X_{t+1}) \sim N(mX_t,m^2p_a + q) \]
Forecast Cycle

- **Forecast Step:**
  \[ P(X_{t+1}) \sim N(\mu_f = mX_t, \quad p_f = m^2p_a + q) \]

- **Analysis Step**
  \[ P(X_{t+1} \mid Y_{t+1}) \sim N(\mu_a, p_a) \]
  \[ 1/p_a = n/p_f + 1/r \]
  \[ \mu_a = (\mu_f/p_f + nY/r) \cdot p_a \]

- Has an analytical solution!

- Kalman Filter

Rudolf Kalman
“Data assimilation isn’t rocket science, but you can use it for that.”

– DAVE MOORE
Generalized to Multivariate

- \((n \times 1)\) vector of state means, \(\mu_a\) or \(\mu_f\)
- \((n \times n)\) state error covariance matrix, \(P_a\) or \(P_f\)  (was \(p_a, p_f\))
- \((p \times 1)\) vector of observations, \(Y\)
- \((p \times p)\) observation error covariance matrix, \(R\)  (was \(r\))
- \((p \times n)\) observation matrix, \(H\)
- \((n \times n)\) linear process model, \(M\)  (was \(m\))
- \((n \times n)\) process error covariance matrix, \(Q\)  (was \(q\))

\[
X_a | Y \sim N(Y | HX_a, R) \quad N(X_a | \mu_f, P_f)
\]
\[ X_a|Y \sim N \left( Y|HX_a, R \right) N \left( X_a|\mu_f, P_f \right) \]

- Solves to be

\[
X_a|Y \sim N \left( \left( H^T R^{-1} H + P_f^{-1} \right)^{-1} \left( H^T R^{-1} Y + P_f^{-1} \mu_f \right), \left( H^T R^{-1} H + P_f^{-1} \right)^{-1} \right)
\]

\[ P_a^{-1} = H^T R^{-1} H + P_f^{-1} \]
\[ X_a | Y \sim N \left( Y | HX_a , R \right) N \left( X_a | \mu_f , P_f \right) \]

- Solves to be
\[ X_a | Y \sim N \left( (H^T R^{-1} H + P_f^{-1})^{-1} (H^T R^{-1} Y + P_f^{-1} \mu_f) , (H^T R^{-1} H + P_f^{-1})^{-1} \right) \]

- Mean and variance simplify to
\[ E \left[ X_a | Y \right] = \mu_a = \mu_f + K \left( Y - H \mu_f \right) \]
\[ Var \left[ X_a | Y \right] = P_a = \left( I - KH \right) P_f \]
\[ K = P_{fH}^T \left( R + H P_f H^T \right)^{-1} \]
Kalman Gain
Example

- Assume $\mu_f = \{\mu_1, \mu_2, \mu_3\}$, $Y = \{y_2, y_3\}$, and observation error is $R = \sigma^2 I$

$$H = \begin{bmatrix} X_1 & X_2 & X_3 \end{bmatrix} \begin{bmatrix} Y_2 \\ Y_3 \end{bmatrix}$$

- The posterior mean for the unobserved $X_1$ is

$$E[x_1] = \mu_1 + w_{12}(y_2 - \mu_2) + w_{13}(y_3 - \mu_3)$$

$$w_1^T = (p_{12} \quad p_{13}) \begin{pmatrix} p_{22} + \sigma^2 & p_{23} \\ p_{32} & p_{33} + \sigma^2 \end{pmatrix}^{-1}$$

If X's are locations and P is a spatial covariance matrix, model is equivalent to Kriging.
Forecast Step \[ X_{t+1} = MX_t + \epsilon \]

The posterior distribution of \( X_{t+1} \) given \( X_t \) is multivariate normal with

\[
\mu_{f,t+1} = E \left[ X_{f,t+1} \mid X_{a,t} \right] = M_t \mu_{a,t}
\]

\[
P_{f,t+1} = Var \left[ X_{f,t+1} \mid X_{a,t} \right] = Q_t + M_t P_{a,t-1} M_t^T
\]
Pro/Con of Kalman Filter (KF)

- Analytically tractable
- Depends only upon the PREVIOUS state, the current Forecast, and the current Data

- Linear
- Normal
- Matrix inversion
- Assumes all parameters (H, R, M, Q) are known
- Forward only
## Uncertainty Propagation

Applied in the Forecast Step

<table>
<thead>
<tr>
<th>Approach</th>
<th>Distribution</th>
<th>Output Moments</th>
<th>Method</th>
</tr>
</thead>
<tbody>
<tr>
<td>Analytic</td>
<td>Variable Transform</td>
<td>Analytical Moments</td>
<td>KF</td>
</tr>
<tr>
<td>Numeric</td>
<td>Monte Carlo</td>
<td>Taylor Series</td>
<td>EKF</td>
</tr>
<tr>
<td></td>
<td>PF</td>
<td>Ensemble</td>
<td>EnKF</td>
</tr>
</tbody>
</table>
Extended Kalman Filter (EKF)

- Addresses **linear** assumption of the Forecast
  
  - $\mu_f = f(\mu_a)$

- Update variance using a Taylor Series expansion
  
  - $F = \text{Jacobian } \left( \frac{df_i}{dx_j} \right)$

  - $P_f \approx Q + F \, P_a \, F^T$  \(\text{(was } Q + M \, P_a \, M^T)\)

- Can be extended to higher orders

- Jensen’s Inequality: Biased, Normality assumption FALSE
\[ N_{t+1} = N_t + rN_t \left(1 + \frac{N}{K}\right) \]

\[ \frac{\partial N_{t+1}}{\partial N_t} = 1 + r - \frac{2r}{K}N_t \]