DART Tutorial Section 1: Filtering For a One Variable System
Example: Estimating the Temperature Outside

An observation has a value (*),

Observed Temperature is 1°C
An observation has a value (\( \ast \)), and an error distribution (red curve) that is associated with the instrument.
Example: Estimating the Temperature Outside

Thermometer outside measures 1°C.

Instrument builder says thermometer is unbiased with +/- 0.8°C Gaussian error.
Thermometer outside measures 1°C.

The red plot is $P(T \mid T_0)$; probability of temperature given that $T_0$ was observed.
Example: Estimating the Temperature Outside

We also have a prior estimate of temperature.

The green curve is $P(T / C)$; probability of temperature given all available prior information $C$. 
Prior information $C$ can include:

1. Observations of things besides $T$;
2. Model forecast made using observations at earlier times;
3. *a priori* physical constraints ($T > -273.15^\circ C$);
4. Climatological constraints ($-30^\circ C < T < 40^\circ C$).
Bayes Theorem: $P(T \mid T_o, C) = \frac{P(T_o \mid T, C)P(T \mid C)}{P(T_o \mid C)}$

**Likelihood**: Probability that $T_o$ is observed if $T$ is true value and given prior information $C$.

**Posterior**: Probability of $T$ given observations and Prior. Also called **update** or **analysis**.
Combining the Prior Estimate and Observation

Rewrite Bayes as:

\[
\frac{P(T_o \mid T,C)P(T \mid C)}{P(T_o \mid C)} = \frac{P(T_o \mid T,C)P(T \mid C)}{\int P(T_o \mid x)P(x \mid C)dx}
\]

\[
= \frac{P(T_o \mid T,C)P(T \mid C)}{\text{normalization}}
\]

Denominator normalizes so Posterior is PDF.
Combining the Prior Estimate and Observation

\[ P(T \mid T_0, C) = \frac{P(T_0 \mid T, C)P(T \mid C)}{\text{normalization}} \]
Combining the Prior Estimate and Observation

\[ P(T | T_0, C) = \frac{P(T_0 | T, C) P(T | C)}{normalization} \]
Combining the Prior Estimate and Observation

\[ P(T \mid T_0, C) = \frac{P(T_0 \mid T, C) P(T \mid C)}{\text{normalization}} \]
Combining the Prior Estimate and Observation

\[ P(T | T_0, C) = \frac{P(T_0 | T, C)P(T | C)}{\text{normalization}} \]
Combining the Prior Estimate and Observation

\[ P(T | T_0, C) = \frac{P(T_0 | T, C)P(T | C)}{\text{normalization}} \]
Bayes’ Rule

\[ p(A \mid BC) = \frac{p(B \mid AC)p(A \mid C)}{p(B \mid C)} = \int p(B \mid x)p(x \mid C) \, dx \]

\( A \): Prior Estimate based on all previous information, \( C \).

\( B \): An additional observation.

\( p(A \mid BC) \): Posterior (updated estimate) based on \( C \) and \( B \).
Bayes’ Rule

\[
p(A \mid BC) = \frac{p(B \mid AC)p(A \mid C)}{p(B \mid C)} = \frac{p(B \mid AC)p(A \mid C)}{\int p(B \mid x)p(x \mid C)\,dx}
\]

\(A\) : Prior Estimate based on all previous information, \(C\).
\(B\) : An additional observation.
\(p(A/BC)\) : Posterior (updated estimate) based on \(C\) and \(B\).
Bayes’ Rule

\[ p(A | BC) = \frac{p(B | AC) p(A | C)}{p(B | C)} = \frac{p(B | AC) p(A | C)}{\int p(B | x) p(x | C) \, dx} \]

- **A**: Prior Estimate based on all previous information, C.
- **B**: An additional observation.
- **p(A/BC)**: Posterior (updated estimate) based on C and B.
Bayes’ Rule

\[ p(A \mid BC) = \frac{p(B \mid AC)p(A \mid C)}{p(B \mid C)} = \frac{p(B \mid AC)p(A \mid C)}{\int p(B \mid x)p(x \mid C) \, dx} \]

**A**: Prior Estimate based on all previous information, \( C \).

**B**: An additional observation.

**\( p(A/BC) \)**: Posterior (updated estimate) based on \( C \) and \( B \).
Bayes’ Rule

\[ p(A \mid BC) = \frac{p(B \mid AC)p(A \mid C)}{p(B \mid C)} = \frac{p(B \mid AC)p(A \mid C)}{\int p(B \mid x)p(x \mid C)\,dx} \]

- **A**: Prior Estimate based on all previous information, \( C \).
- **B**: An additional observation.
- **\( p(A \mid BC) \)**: Posterior (updated estimate) based on \( C \) and \( B \).
Color Scheme

Green == Prior

Red == Observation

Blue == Posterior

The same color scheme is used throughout ALL Tutorial materials.
Product of Two Gaussians

\[ p(A \mid BC) = \frac{p(B \mid AC)p(A \mid C)}{p(B \mid C)} = \frac{p(B \mid AC)p(A \mid C)}{\int p(B \mid x)p(x \mid C) \, dx} \]

Any 1-D normal distribution can be represented as a PDF:

\[ \frac{1}{\sigma_x \sqrt{2\pi}} \exp\left( -\frac{(x - \mu_x)^2}{2\sigma_x^2} \right) \]
This product is closed for Gaussian distributions.

\[
p(A \mid BC) = \frac{p(B \mid AC) p(A \mid C)}{p(B \mid C)} = \int p(B \mid x) p(x \mid C) \, dx
\]
Product of d-dimensional normals with means $\mu_1$ and $\mu_2$ and covariance matrices $\Sigma_1$ and $\Sigma_2$ is normal.

$$N(\mu_1, \Sigma_1)N(\mu_2, \Sigma_2) = cN(\mu, \Sigma)$$

Covariance: $$\Sigma = (\Sigma_1^{-1} + \Sigma_2^{-1})^{-1}$$

Mean: $$\mu = (\Sigma_1^{-1} + \Sigma_2^{-1})^{-1}(\Sigma_1^{-1}\mu_1 + \Sigma_2^{-1}\mu_2)$$

Weight: $$c = \frac{1}{(2\pi)^{d/2}|\Sigma_1 + \Sigma_2|^{1/2}} \exp\left\{-\frac{1}{2}\left[(\mu_2 - \mu_1)^T(\Sigma_1 + \Sigma_2)^{-1}(\mu_2 - \mu_1)\right]\right\}$$

The weight is simply the normalization of the normal distribution defined by the product of the prior and observation likelihood.
Product of Two Gaussians

\[
p(A \mid BC) = \frac{p(B \mid AC)p(A \mid C)}{p(B \mid C)} = \frac{p(B \mid AC)p(A \mid C)}{\int p(B \mid x)p(x \mid C)\,dx}
\]

This product can be determined analytically for Gaussian distributions.

But, for general distributions, there’s no analytical product.
Don’t know much about properties of this sample. May naively assume it is random draw from ‘truth’.

Ensemble filters: Prior is available as finite sample.
Product of Two Gaussians

$$p(A \mid BC) = \frac{p(B \mid AC)p(A \mid C)}{p(B \mid C)} = \int p(B \mid x)p(x \mid C) \, dx$$

How can we take product of sample with continuous likelihood?

Fit a continuous (Gaussian for now) distribution to sample.
Product of Two Gaussians

\[
p(A \mid BC) = \frac{p(B \mid AC)p(A \mid C)}{p(B \mid C)} = \int p(B \mid x)p(x \mid C) \, dx
\]

Observation likelihood usually continuous (nearly always Gaussian).
Product of prior Gaussian fit and Obs. likelihood is Gaussian.

\[ p(A \mid BC) = \frac{p(B \mid AC)p(A \mid C)}{p(B \mid C)} = \frac{p(B \mid AC)p(A \mid C)}{\int p(B \mid x)p(x \mid C)dx} \]

Product of Two Gaussians

Analytically computing continuous posterior is simple. BUT, we need to have a SAMPLE of this PDF...
There are many ways to do this.

Exact properties of different methods may be unclear. Trial and error still best way to see how they perform. Will interact with properties of prediction models, etc.
Ensemble Adjustment (Kalman) Filter

Sampling Posterior PDF

Ensemble Adjustment (Kalman) Filter

Prior Ensemble

---

Probability

-4
-2
0
2
4

0
0.2
0.4
0.6
Again, fit a Gaussian to sample.
Compute posterior PDF (same as previous algorithms).
Use deterministic algorithm to ‘adjust’ ensemble.
1. ‘Shift’ ensemble to have exact mean of posterior.
2. Use linear contraction to have exact variance of posterior.
$x_i^u = \left( x_i^p - \bar{x}^p \right) \cdot \left( \sigma^u / \sigma^p \right) + \bar{x}^u$

$p$ is prior,
$u$ is update (posterior),
$\sigma$ is standard deviation,
overbar is ensemble mean.
Bimodality maintained, but not appropriately positioned or weighted. No problem with random outliers.