Gaussian Filtering Strategies for Nonlinear Systems
Canonical Nonlinear Filtering Problem

\[ \tilde{u}_{m+1} = \tilde{f}(\tilde{u}_m) + \tilde{\sigma}_{m+1} \]
\[ \tilde{v}_{m+1} = \tilde{g}(\tilde{u}_{m+1}) + \tilde{\sigma}^o_{m+1} \]

- \( \tilde{f} \) and \( \tilde{g} \) are nonlinear & deterministic
- Noise/Errors \( \tilde{\sigma} \) are Gaussian with zero mean
- Continuous-time dynamics can be written in this form
- Observation error has covariance \( R_o \) (I use subscript \( o \); book uses superscript)
- In general the optimal Bayesian filter is not the Kalman filter because the distributions are not Gaussian.

This lecture introduces a range of Gaussian filtering strategies for nonlinear systems.
OI and 3DVAR

Suppose that you are given a prior mean $\tilde{u}_{m+1|m}$ and covariance $R_{m+1|m}$, and an observation $\tilde{v}_{m+1}$ with an observation error covariance $R_0$.

Once these are given, the maximum likelihood state for the Bayesian posterior minimizes the following nonlinear functional

$$J(\tilde{u}) = (\tilde{u} - \tilde{u}_{m+1|m})^T R_{m+1|m}^{-1} (\tilde{u} - \tilde{u}_{m+1|m}) +$$

$$(\tilde{v}_{m+1} - \tilde{g}(\tilde{u}))^T R_0^{-1} (\tilde{v}_{m+1} - \tilde{g}(\tilde{u})).$$

3DVAR (3D variational) methods solve this minimization problem, typically using an iterative method. They typically specify a time-independent background covariance matrix $R_{m+1|m} = B$ that has a known eigenvalue decomposition, and the prior mean is given by one simulation (or an average over many) of the nonlinear dynamics.
Optimal Interpolation (OI) also specifies a time-independent background covariance matrix $R_{m+1|m} = B$, but relies on a linear/linearized observation operator and solves the Kalman filter equations for the posterior mean:

$$
\tilde{\mathbf{u}}_{m+1|m+1} = \tilde{\mathbf{u}}_{m+1|m} + \mathbf{K}(\tilde{\mathbf{v}}_{m+1} - \mathbf{G}\tilde{\mathbf{u}}_{m+1|m})
$$

$$
\mathbf{K} = \mathbf{B}\mathbf{G}^T(\mathbf{G}\mathbf{B}\mathbf{G}^T + \mathbf{R}_o)^{-1}
$$

The prior mean can be specified either by a single simulation of the nonlinear dynamics, or the climatological mean state.

Some “observational” data products (like AVISO ocean SSH data) are actually raw satellite data processed using OI.
OI and 3DVAR

If the background covariance and prior mean are the same, then OI and 3DVAR are mathematically equivalent.

The background covariance matrix is meant to represent forecast uncertainty, and specifying it is something of a black art.

Operational weather forecasting and reanalysis moved away from 3DVAR in the late 90s, but 3DVAR and OI are still found in some ocean data assimilation systems.

One might hope for better results by using a time-dependent prior covariance, and the next strategies do this.
Extended Kalman Filter

The EKF linearizes the dynamic and observation operators about the current mean state:

\[
\tilde{f}(\tilde{u}) \approx \tilde{f}(\tilde{u}_{m|m}) + F_m \cdot (\tilde{u} - \tilde{u}_{m|m}) \\
\tilde{g}(\tilde{u}) \approx \tilde{g}(\tilde{u}_{m|m}) + G_m \cdot (\tilde{u} - \tilde{u}_{m|m})
\]

Linearization is accurate only for small deviations about the mean state.

The linearized system keeps distributions Gaussian, so we can use the Kalman filter to update the mean and covariance.
Extended Kalman Filter

The EKF can be

- inaccurate because the deviations from the mean are not small
- impractical because the nonlinear operators can be difficult or impossible to linearize
- impractical because the system dimension is so large that it’s impossible to implement

For these reasons EKF is seldom used in atmosphere/ocean applications.
Ensemble Kalman Filter

The essential idea of the Ensemble Kalman filter (orig. Evensen 1994) is to approximate the mean and covariance using an ensemble with $K$ members$^1$.

$$\vec{u} \approx \frac{1}{K} \sum_{k=1}^{K} \vec{u}^k$$

$$R_{m+1|m} \approx \frac{1}{K-1} \sum_{k=1}^{K} (\vec{u}^k - \vec{u})(\vec{u}^k - \vec{u})^T$$

$$= \frac{1}{K-1} U_{m+1|m} U_{m+1|m}^T$$

The columns of the matrix $U_{m+1|m}$ are the ensemble perturbations $\vec{u}^k - \vec{u}$.

The ensemble covariance matrix has rank at most $K-1$.

$^1$Try not to confuse the ensemble size $K$ with the Kalman gain matrix $K$.
Ensemble Kalman Filter

The ensemble Kalman gain matrix has the form

\[ K = R_{m+1|m} G^T \left( G R_{m+1|m} G^T + R_o \right)^{-1} \]

\[ \approx U_{m+1|m} \left[ \left( G U_{m+1|m} \right)^T \left( \left( G U_{m+1|m} \right) \left( G U_{m+1|m} \right)^T + (K - 1)R_o \right)^{-1} \right] \]

\[ = U_{m+1|m} \left[ V^T \left( V V^T + (K - 1)R_o \right)^{-1} \right]. \]

For nonlinear observation operators define the columns of \( V \) to be \( \tilde{g}(\tilde{u}^k) - \tilde{g}(\tilde{u}) \) so that \( V V^T / (K - 1) \) estimates the covariance matrix in observation space.
Ensemble Kalman Filter

The ensemble Kalman gain matrix has the form

\[
K = U_{m+1|m} \left[ V^T \left( VV^T + (K - 1)R_o \right)^{-1} \right].
\]

Since the KF update of the mean is of the form

\[
\tilde{u}_{m+1|m+1} = \tilde{u}_{m+1|m} + K(\tilde{v}_{m+1} - \tilde{g}(\tilde{u}_{m+1|m}))
\]

the ensemble KF update to the mean is in the range of \( U_{m+1|m} \). This can be important if the ensemble perturbations do not span the space.
Ensemble Kalman Filter

We have seen how to update the mean in the ensemble Kalman filter. How do you update the ensemble members?

The original EnKF (Evensen 1994) proposed to use the formula for the mean to update each ensemble member, i.e.

$$\tilde{u}^{k}_{m+1|m+1} = \tilde{u}^{k}_{m+1|m} + K(\tilde{v}_{m+1} - \tilde{g}(\tilde{u}^{k}_{m+1|m}))$$

**Exercise:** Show that the posterior ensemble covariance does not match the posterior covariance from the Kalman filter formulas.
Ensemble Kalman Filter

The solution is to generate an ensemble of observations by sampling from the observation error distribution:

\[
\tilde{v}_{m+1}^k = \tilde{v}_{m+1} + \tilde{\sigma}_m^o, \quad \tilde{u}_{m+1|m+1}^k = \tilde{u}_{m+1|m}^k + K(\tilde{v}_{m+1}^k - \tilde{g}(\tilde{u}_{m+1|m}^k))
\]

This is called “the” EnKF or the “stochastic” or “perturbed obs” EnKF.

The posterior perturbations are all in the range of the prior perturbations.
Ensemble Square Root Filter

- Use ensemble to estimate the statistics of state; nonlinear dynamics
- Change ensemble to exactly match the posterior mean and covariance
- Serial assimilation
  Possible for observation batches uncorrelated to each other
  No big matrix calculation
  Easy to implement
Univariate assimilation: EAKF for scalar model

\[
\begin{align*}
\text{state variable} & \quad u_{m+1} &= f(u_m) + \sigma_m \\
\text{obs} & \quad v_{m+1} &= g(u_{m+1}) + \sigma^o = u_{m+1} + \sigma^o
\end{align*}
\]

(1)

\[u_{m+1}, v_{m+1} \in \mathbb{R}^1, \text{ and } (\sigma^o)^2 = r^o.\] General observation operator \(g\) will be considered later.
Use sample mean $\bar{u}_f = \frac{1}{K} \sum_{k=1}^{K} u_k$ and sample variance $r_f = \frac{1}{K - 1} \sum_{k=1}^{K} (u_k - \bar{u})^2$ to determine the prior distribution PDF.

Fit a Gaussian to the sample $\mathcal{N}(\bar{u}_f, r_f)$ at $t = t_2$. 
Get the observation likelihood $\mathcal{N}(v, r^\circ)$. 
Compute the continuous posterior PDF. Posterior mean $\bar{u}^a$ and variance $r^a$ by 1D Kalman formula

\begin{align*}
\bar{u}^a &= \bar{u}^f + \frac{r_f}{r_f + r^o}(v - \bar{u}^f) \\
r^a &= \left(\frac{1}{r_f} + \frac{1}{r^o}\right)^{-1}
\end{align*}  \tag{2}
Use a deterministic algorithm to adjust the ensemble.

1. Shift the ensemble to have the exact posterior mean $\bar{u}^a$.
2. Linearly contract the ensemble to have the exact posterior variance $r^a$. 
Ensemble Square Root Kalman Filters

The stochastic ensemble Kalman filter generates a posterior ensemble such that the expected value of the posterior ensemble covariance matches the true Kalman filter posterior covariance.

Ensemble “Square Root” Kalman filters were developed such that the posterior ensemble covariance exactly matches the true Kalman filter posterior covariance.

There are several kinds of square root EnKFs; we focus on two defined as follows:

ETKF: $U_{m+1|m+1} = U_{m+1|m} T$

EAKF: $U_{m+1|m+1} = AU_{m+1|m}$

Note: the sum of the ensemble perturbations should be zero by definition, i.e. $U \cdot 1 = 0$ where $1$ is a vector of ones.

For the ETKF this means that $1$ must be an eigenvector of $T$. 
Ensemble Square Root Kalman Filters

The goal is to find matrices \( T \) or \( A \) such that the posterior covariance satisfies the Kalman equation

\[
R_{m+1|m+1} = (I - KG) R_{m+1|m}.
\]

Written in terms of the ensemble perturbation matrices this is

\[
U_{m+1|m+1} U_{m+1|m+1}^T = (I - KG) U_{m+1|m} U_{m+1|m}^T
\]

Recalling the ensemble definition of the Kalman gain, this is

\[
U_{m+1|m+1} U_{m+1|m+1}^T = U_{m+1|m} \left[ I - V^T (V V^T + (K - 1) R_o)^{-1} V \right] U_{m+1|m}^T
\]

\( T \) is a square root of the matrix in square brackets.
ETKF

Eigenvalue decomposition: $Z$ is real orthogonal and $\Lambda$ is diagonal

$$(VV^T + (K - 1)R_o)^{-1} = Z^T \Lambda^{-1} Z$$

$$I - V^T(VV^T + (K - 1)R_o)^{-1}V = I - \left(\Lambda^{-1/2}ZV\right)^T \left(\Lambda^{-1/2}ZV\right)$$

Now SVD: $S$ and $W$ real orthogonal and $D$ diagonal

$$\left(\Lambda^{-1/2}ZV\right) = WDS$$

$$I - \left(\Lambda^{-1/2}ZV\right)^T \left(\Lambda^{-1/2}ZV\right) = I - S^T D^2 S = S^T (I - D^2) S$$

So $S^T (I - D^2)^{1/2}$ is a candidate for the ETKF transformation matrix $T$. Sakov & Oke (MWR 2008) show that the diagonal matrix $I - D^2$ is positive definite, so the real square root exists.
ETKF

$S^T(I - D^2)^{1/2}$ is a candidate for the ETKF transformation matrix $T$, but it does not preserve zero mean perturbations, i.e. 1 is not an eigenvalue.

Sakov & Oke (2008) and Livings et al. (2008) showed that the “symmetric” square root transformation matrix

$S^T(I - D^2)^{1/2}S$

does preserve the mean.

It may also be advantageous in some situations to add a random rotation of the ensemble by appending a mean-preserving random orthogonal matrix $\Theta$ to the transformation

$S^T(I - D^2)^{1/2}S\Theta$

Sakov & Oke (2008) show how to construct such a $\Theta$. 
I will state the formula for an EAKF adjustment matrix $A$, and then prove that it gives the same posterior covariance as the ETKF.

Let the following be a scaled eigenvalue decomposition of the prior covariance matrix

$$U_{m+1|m}U^T_{m+1|m} = F\Sigma^2 F^T$$

where $F$ is real orthogonal and $\Sigma^2$ is diagonal.

The adjustment matrix is

$$A = U_{m+1|m}S^T(I - D^2)^{1/2}\Sigma^{-1}F^T$$
Plugging the EAKF adjustment matrix in to the covariance update formula gives

\[ U_{m+1|m+1} U^T_{m+1|m+1} = A U_{m+1|m} U^T_{m+1|m} A^T \]

\[ = \left( U_{m+1|m} S^T (I - D^2)^{1/2} \right) \Sigma^{-1} F^T \Sigma^2 F^T \Sigma^{-1} \]

\[ \left( U_{m+1|m} S^T (I - D^2)^{1/2} \right)^T \]

\[ = U_{m+1|m} TT^T U^T_{m+1|m} \]

The posterior covariance is equal to the ETKF posterior covariance, which matches the Kalman posterior covariance exactly.
Note that although this gives the same posterior covariance as the ETKF, the posterior ensemble is not the same. The posterior ensembles are related by

ETKF: \( U_{m+1|m+1} = U_{m+1|m} S^T (I - D^2)^{1/2} S \)

EAKF: \( U_{m+1|m+1} = U_{m+1|m} S^T (I - D^2)^{1/2} \Sigma^{-1} F^T U_{m+1|m} \)

The ensembles differ only by an orthogonal rotation.

Note that for both ETKF and EAKF the posterior ensemble perturbations are in the space spanned by the prior perturbation ensemble.
Ensemble Square Root Filters

The foregoing development of the ETKF follows Evensen’s book “Data Assimilation”, available from the NYU library.

Efficient implementations of ensemble square root filters will be discussed next lecture.

There are other square root filters not covered here.

Sometimes the square root filters are called “deterministic” EnKF, but the use of random rotations in ETKF means this is not a great name.
Covariance Inflation

In ensemble Kalman filters the prior ensemble covariance is often too small, e.g. because the ensemble does not sample well and/or because the model has errors.

One *ad hoc* way to remedy this is to inflate the prior ensemble covariance, either by adding random perturbations (realizations of model error) to the ensemble members or by multiplicatively increasing the ensemble perturbations.

We will only use multiplicative inflation, which takes the form

\[ U_{m+1|m} \leftarrow \sqrt{1 + rU_{m+1|m}} \]

where \( r > 0 \) is an empirically tuned constant. This has the effect of causing the filter to weight the observations more heavily.
- Increase the spread in the prior.
- Give more weight to the observation, less to prior.

⇒ Recover practical controllability
Inflate all state variables by same amount before assimilation

Capabilities:
1. Can maintain linear balances.
2. Stays on local flat manifolds.
3. Simple and inexpensive.

Liabilities:
- State variables not constrained by observations can blow up!
- Inflation magnitude is selected by trial and error.
Finite samples estimate low correlations less accurately than high correlations. The above plot shows the error variance in $10^4$ repeated estimates of the correlation between two standard Gaussian variables using $N$ ensemble members.
Covariance Localization

Variables that are far from each other often have low correlations, but these low correlations are poorly estimated by finite ensembles.

One *ad hoc* way to remedy this is to scale the prior covariance matrix as follows

\[ R_{m+1|m} \leftarrow \rho \circ R_{m+1|m} \]

where \( \rho \) is a sparse correlation matrix and \( \circ \) denotes elementwise multiplication (aka Schur or Hadamard product, depending on whether you’re German or French).

The correlation matrix can be specified in a variety of ways, some theory and examples are in Gaspari & Cohn, “Construction of correlation functions in two and three dimensions” QJRMS 1999.

The Schur Product Theorem guarantees that the elementwise product of two covariance matrices is a covariance matrix (e.g. “Matrix Analysis”, Horn & Johnson, Cambridge U Press).
Covariance Localization

**Example:** Consider a 5 component system on a periodic domain with one observation of the third component. The Kalman filter update is

\[
\vec{u}_{m+1|m+1} = \vec{u}_{m+1|m} + \frac{(v - \bar{u}_3)}{r_{3,3} + r_o} \begin{pmatrix} r_{1,3} \\ r_{2,3} \\ r_{3,3} \\ r_{4,3} \\ r_{5,3} \end{pmatrix}
\]

were \( r_{i,j} \) are elements of the prior covariance \( R_{m+1|m} \).

All 5 variables are updated based on an observation of only one variable.
Covariance Localization

If localization is used with a correlation matrix of the form

\[ \rho = \begin{pmatrix}
1 & 1/2 & 0 & 0 & 1/2 \\
1/2 & 1 & 1/2 & 0 & 0 \\
0 & 1/2 & 1 & 1/2 & 0 \\
0 & 0 & 1/2 & 1 & 1/2 \\
1/2 & 0 & 0 & 1/2 & 1
\end{pmatrix} \]

The Kalman mean update becomes

\[ \tilde{u}_{m+1|m+1} = \tilde{u}_{m+1|m} + \frac{(v - \bar{u}_3)}{r_{3,3} + r_o} \begin{pmatrix}
0 \\
r_{2,3}/2 \\
r_{3,3} \\
r_{4,3}/2 \\
0
\end{pmatrix} \]

Localization guarantees that one observation only impacts variables at nearby locations.
How to remedy regression sampling error?

- Use larger ensembles; but expensive for large models
- Use additional a priori information about relation between observations and state variables. Do not let an observation impact a state variable if they are known to be unrelated.

**Localization**: try reducing regression factor as function of distance between observation and state variable. Compactly supported fifth order polynomial by Gaspari-Cohn is most commonly used for geophysics.
Covariance Localization

The ensemble prior covariance can be rank-deficient. Localization typically increases the rank of the prior covariance matrix; e.g. extreme localization means the prior covariance is full rank. This means that the analysis updates are not restricted to the range of the prior ensemble perturbation matrix (a good thing).

It also means that the posterior ensemble perturbation matrix is no longer guaranteed to match the Kalman posterior covariance (possibly not good).

There are other kinds of localization, e.g. LETKF.

Localization can lead to imbalance (the generation of fast waves).
Pseudocode of serial EAKF for one assimilation cycle

Inflate covariance;

for each uncorrelated observation do
  Calculate increments of ensemble;
  for other state variables do
    Find linear regression fit;
    Reduce regression factor if necessary (Localization);
    Update the state variables of each ensemble member;
  end
end