Basic Equations for the Ensemble Kalman and Square-root Filters

List of symbols

\( \psi^f \)  Forecast state (ensemble member \( i \))
\( \psi^t \)  Truth state
\( \psi^a_i \)  Analysis state (ensemble member \( i \))
\( y_i \)  Observation set (associated with ensemble member \( i \))
\( y_i^{(m)} \)  Model observation set (associated with ensemble member \( i \))
\( P^f \)  Forecast error covariance matrix
\( P^f_e \)  Ensemble estimate of forecast error covariance matrix
\( P^a \)  Analysis error covariance matrix
\( P^a_e \)  Ensemble estimate of analysis error covariance matrix
\( R \)  Observation error covariance matrix
\( h \)  Non-linear observation operator
\( H \)  Tangent linear model of observation operator
\( K \)  Kalman gain matrix
\( A^f \)  Forecast state vector matrix (forecasts are columns)
\( A^f_s \)  Forecast state vector perturbation matrix
\( A^a \)  Analysis state vector matrix (analyses are columns)
\( A^a_s \)  Analysis state vector perturbation matrix
\( D \)  Observation vector matrix (observation sets are columns)
\( \Upsilon \)  Observation vector perturbation matrix
\( I_{N \times N} \)  \( N \times N \) identity matrix
\( 1_{N \times N} \)  \( N \times N \) matrix of \( 1/N \) values
\( S \)  Model observation perturbations calculated from forecast perturbations
\( C \)  \((N - 1) \times \) model ob + ob covariance matrices

Let there be \( n \) elements in a state vector, \( N \) ensemble members and \( p \) observations.
Definition of error covariance matrices

Forecast error covariance matrices
\[
\begin{align*}
P_f &= (\psi_f' - \overline{\psi_f'})(\psi_f' - \overline{\psi_f'})^T \\
P_{ef} &= (\psi_f' - \overline{\psi_e'})(\psi_f' - \overline{\psi_e'})^T
\end{align*}
\]

Analysis error covariance matrices
\[
\begin{align*}
P_a &= (\psi_a' - \psi_f')(\psi_a' - \psi_f')^T \\
P_{ea} &= (\psi_a' - \psi_a')(\psi_a' - \psi_a')^T
\end{align*}
\]

Observation error covariance matrix
\[
R_e = (y_i - \overline{y})(y_i - \overline{y})^T
\]

The observation operator
\[
y_i^{(m)} = \mathbf{h}(\overline{\psi}^i) + \mathbf{H}(\psi_f' - \overline{\psi_f'})
\]

The KF analysis update step for each member and for the mean

Analysis step for ensemble member \(i\)
\[
\begin{align*}
\psi_a^i &= \psi_f^i + P_e^i H^T (H P_e^i H^T + R_e)^{-1} (y_i - \mathbf{h}(\psi_f^i) - H(\psi_f^i - \overline{\psi_f'})) \\
\overline{\psi_a^i} &= \overline{\psi_f^i} + P_e^i H^T (H P_e^i H^T + R_e)^{-1} (\overline{y}_i - \mathbf{h}(\overline{\psi_f'}))
\end{align*}
\]

Deviation of each member from the mean
\[
\begin{align*}
\psi_a^i - \overline{\psi_a^i} &= \psi_f^i - \overline{\psi_f'} + P_e^i H^T (H P_e^i H^T + R_e)^{-1} (y_i - \overline{y}_i - H(\psi_f^i - \overline{\psi_f'})) \\
&= (I - KH)(\psi_f^i - \overline{\psi_f'}) + K(y_i - \overline{y}_i) \\
K &= P_e^i H^T (H P_e^i H^T + R_e)^{-1}
\end{align*}
\]

The ensemble analysis error covariance

Assume no correlation between forecast and observation errors.
\[
\begin{align*}
P_{ea} &= \left\langle (I - KH)(\psi_f^i - \overline{\psi_f'}) + K(y_i - \overline{y}_i) \left( (I - KH)(\psi_f^i - \overline{\psi_f'}) + K(y_i - \overline{y}_i) \right)^T \right\rangle \\
&= \left\langle (I - KH)(\psi_f^i - \overline{\psi_f'}) \left( (I - KH)(\psi_f^i - \overline{\psi_f'}) \right)^T \right\rangle + \left\langle K(y_i - \overline{y}_i) \left( K(y_i - \overline{y}_i) \right)^T \right\rangle \\
&= (I - KH) P_e^i (I - KH)^T + KRK_T^T \\
&= P_e^i - P_e^i H^T K_T^T - KHP_e^i + KHP_e^i H^T K_T^T + KRK_T^T \\
&= P_e^i + K(H P_e^i H^T + R_e) K_T^T - P_e^i H^T K_T^T - KHP_e^i \\
&= P_e^i + P_e^i H^T (H P_e^i H^T + R_e)^{-1} (H P_e^i H^T + R_e) K_T^T - P_e^i H^T K_T^T - KHP_e^i
\end{align*}
\]
\[
\begin{align*}
\mathbf{P}_e' &= \mathbf{K} \mathbf{H} \mathbf{P}_e' \\
&= (1 - \mathbf{K}) \mathbf{P}_e'
\end{align*}
\]

Matrix representation of the ensemble

The forecast states \( \psi_f' \) may be assembled into columns of the \( n \times N \) matrix \( \mathbf{A}' \).
\[
\mathbf{A}' = (\psi_1' \psi_2' \ldots \psi_N')
\]
The analysis states \( \psi_a' \) may be assembled into columns of the \( n \times N \) matrix \( \mathbf{A}'' \).
\[
\mathbf{A}'' = (\psi_1'' \psi_2'' \ldots \psi_N'')
\]
The perturbed observations \( \mathbf{y}_i \) may be assembled into the columns of the \( p \times N \) matrix \( \mathbf{D} \).
\[
\mathbf{D} = (\mathbf{y}_1 \mathbf{y}_2 \ldots \mathbf{y}_N)
\]
The ensemble means, \( \bar{\mathbf{A}}' \), \( \bar{\mathbf{A}}'' \) and \( \bar{\mathbf{D}} \) are matrices that comprise the respective ensemble mean state repeated in each column.
\[
\begin{align*}
\bar{\mathbf{A}}' &= \mathbf{A}' \mathbf{1}_{N \times N} \\
\bar{\mathbf{A}}'' &= \mathbf{A}'' \mathbf{1}_{N \times N} \\
\bar{\mathbf{D}} &= \mathbf{D} \mathbf{1}_{N \times N}
\end{align*}
\]
\[
\mathbf{1}_{N \times N} = \begin{pmatrix}
1/N & 1/N & \ldots & 1/N \\
1/N & 1/N & \ldots & 1/N \\
\ldots & \ldots & \ldots & \ldots \\
1/N & 1/N & \ldots & 1/N
\end{pmatrix}
\]

Matrix representation of the ensemble perturbations from the mean

\[
\begin{align*}
\mathbf{A}'' &= \mathbf{A}' - \bar{\mathbf{A}}' = \mathbf{A}' (\mathbf{I}_{N \times N} - \mathbf{1}_{N \times N}) \\
\mathbf{A}'' &= \mathbf{A}'' - \bar{\mathbf{A}}'' = \mathbf{A}'' (\mathbf{I}_{N \times N} - \mathbf{1}_{N \times N}) \\
\mathbf{Y} &= \mathbf{D} - \bar{\mathbf{D}} = \mathbf{D} (\mathbf{I}_{N \times N} - \mathbf{1}_{N \times N})
\end{align*}
\]
Note: \( (\mathbf{I}_{N \times N} - \mathbf{1}_{N \times N})^2 = (\mathbf{I}_{N \times N} - \mathbf{1}_{N \times N}) \)

The error covariance matrices from the matrix representation of ensemble perturbations

\[
\begin{align*}
\mathbf{P}_e' &= \frac{1}{N - 1} \mathbf{A}' \mathbf{A}'^T \\
\mathbf{P}_e'' &= \frac{1}{N - 1} \mathbf{A}'' \mathbf{A}''^T \\
\mathbf{R}_e &= \frac{1}{N - 1} \mathbf{Y} \mathbf{Y}^T
\end{align*}
\]

The analysis equation in terms of the matrix representation of the ensemble

\[
\begin{align*}
\mathbf{A}'' &= \mathbf{A}' + \mathbf{K} (\mathbf{D} - \mathbf{H} \bar{\mathbf{A}}') \\
\mathbf{K} &= \mathbf{P}_e' \mathbf{H}^T (\mathbf{H} \mathbf{P}_e' \mathbf{H}^T + \mathbf{R}_e)^{-1} \\
&= \mathbf{A}' \mathbf{A}'^T \mathbf{H}^T (\mathbf{H} \mathbf{A}' \mathbf{A}'^T \mathbf{H}^T + \mathbf{Y} \mathbf{Y}^T)^{-1} \\
&= \mathbf{A}' (\mathbf{H} \mathbf{A}' \mathbf{A}'^T (\mathbf{H} \mathbf{A}' \mathbf{A}'^T + \mathbf{Y} \mathbf{Y}^T)^{-1}
\end{align*}
\]

These equations do not depend upon explicit knowledge of the forecast error covariance matrix. It is implied by the ensemble of forecast states. Warning: there is a possible rank deficiency if \( N < p \).
The ensemble mean analysis equation in terms of matrix representation of the ensemble

\[
\tilde{A}^a = A^a 1_{N \times N} = A^f 1_{N \times N} + K(D - H\tilde{A}^\prime) 1_{N \times N}
\]

\[
= \tilde{A}^f + K(D - H\tilde{A}^\prime)
\]

The analysis error covariance in terms of matrix representations of the ensemble

\[
P_e^a = \frac{1}{N - 1} A^a A^a^T - \frac{1}{N - 1} A^a (I_{N \times N} - 1_{N \times N}) (I_{N \times N} - 1_{N \times N}) A^a^T
\]

\[
= \frac{1}{N - 1} A^a (I_{N \times N} - 1_{N \times N}) A^a^T
\]

\[
= \frac{1}{N - 1} [A^f + K(D - H\tilde{A}^\prime)] (I_{N \times N} - 1_{N \times N}) [A^f + K(D - H\tilde{A}^\prime)]^T
\]

\[
= P_e^f - P_e^f H^T K - KHP_e^f + KR_e K^T + KHP_e^f H^T K
\]

\[
= (I - KH) P_e^f \quad \text{using analysis beforehand.}
\]

Factorization in terms of \(S\) and \(C\) matrices

The ensemble mean analysis can be written in terms of matrices \(S\) and \(C\).

\[
\tilde{A}^a = \tilde{A}^f + A^a S^T C^{-1} (D - H\tilde{A}^\prime)
\]

where

\[
S = HA^f \quad C = SS^T + (N - 1) R_e
\]

\(S\) is the matrix containing model observation perturbations (calculated from the matrix of forecast perturbations), and \(C\) is \((N - 1)\) times the sum of model observation covariances, \(SS^T\), and observation error covariances, \(R_e\).

The above result can be checked by substitution

\[
\tilde{A}^a = \tilde{A}^f + A^f A^f^T H^T [HA^f A^f^T H^T + (N - 1) R_e]^{-1} (D - H\tilde{A}^\prime)
\]

\[
= \tilde{A}^f + (N - 1) P_e^f H^T [H(N - 1) P_e^f H^T + (N - 1) R_e]^{-1} (D - H\tilde{A}^\prime)
\]

\[
= \tilde{A}^f + P_e^f H^T [HP_e^f H^T + R_e]^{-1} (D - H\tilde{A}^\prime)
\]

\[
= \tilde{A}^f + K(D - H\tilde{A}^\prime) \quad \text{expression is confirmed.}
\]

\(P_e^a\) can also be written in terms of matrices \(S\) and \(C\).

\[
P_e^a = \frac{1}{N - 1} A^a (I - S^T C^{-1} S) A^a^T
\]

This result can also be checked by substitution

\[
P_e^a = \frac{1}{N - 1} A^a A^a^T - \frac{1}{N - 1} A^a A^a^T H^T [HA^f A^f^T H^T + (N - 1) R_e]^{-1} H A^f A^a^T
\]

\[
= P_e^f - P_e^f H^T [H(N - 1) P_e^f H^T + (N - 1) R_e]^{-1} (N - 1) H P_e^f
\]

\[
= P_e^f - P_e^f H^T [HP_e^f H^T + R_e]^{-1} H P_e^f
\]
\[ P'_e = KHP'_e \]
\[ (I - KH)P'_e \]
expression is confirmed.
The expressions involving \( S \) and \( C \) are useful because the matrices \( S \) and \( C \) are calculable for a given ensemble. Both exist in observation space.

**A 'square root' analysis scheme**

It is desirable to calculate the square-root of \( P'_e \). The square-root matrix may be regarded as a set of ensemble perturbations that have covariance \( P'_e \). Given that \( C \) is available, write it in terms of its eigenvalue decomposition

\[ C = Z\Lambda Z^T \]

where \( \Lambda \) is the (diagonal) matrix of eigenvalues and \( Z \) is the orthonormal matrix of eigenvectors. We assume that all eigenvalues of \( C \) are non-zero, which will allow calculation of its inverse. \( P'_e \) is then written

\[ P'_e = \frac{1}{N-1}A'^T(I - S^T Z^{-1} Z^T S) A'^T \]
\[ = \frac{1}{N-1}A'^T(I - X^TX) A'^T \]

where \( X = \Lambda^{-1/2}Z^TS \)

Matrix \( X \) may be further written in terms of a singular value decomposition in the following way

\[ X = U\Sigma V^T \]

where \( \Sigma \) is the matrix of singular values, \( U \) is the orthonormal matrix of left singular vectors and \( V \) is the orthonormal matrix of right singular vectors. \( P'_e \) is then written

\[ P'_e = \frac{1}{N-1}A'^T(I - V\Sigma^TU^T U\Sigma V^T) A'^T \]
\[ = \frac{1}{N-1}A'^T(I - V\Sigma^T V) A'^T \]
\[ = \frac{1}{N-1}A'^T V(I - \Sigma^T \Sigma)^{1/2} V^T A'^T \]

\( \Sigma^T \Sigma \) is a diagonal matrix of size \( N \times N \). The square-root of the combined diagonal matrix \( I - \Sigma^T \Sigma \) is therefore trivial to compute.

\[ P'_e = \frac{1}{N-1}A'^a A'^a^T = \frac{1}{N-1}A'^a V(I - \Sigma^T \Sigma)^{1/2} (I - \Sigma^T \Sigma)^{1/2} V^T A'^a \]
\[ \therefore A'^a A'^a^T = [A'^a V(I - \Sigma^T \Sigma)^{1/2}] [A'^a V(I - \Sigma^T \Sigma)^{1/2}]^T \]
\[ A'^a = A'^a V(I - \Sigma^T \Sigma)^{1/2} \]