Basic Equations for the Ensemble Kalman and Square-root Filters

Ross Bannister, January 2008 / July 2008

List of symbols

ψ_i^f	Forecast state (ensemble member i)
ψ^t	Truth state
ψ_i^a	Analysis state (ensemble member i)
$\boldsymbol{y}_i$	Observation set (associated with ensemble member i)
$\boldsymbol{y}_i^{(m)}$	Model observation set (associated with ensemble member i)
$\mathbf{P}^f$	Forecast error covariance matrix
$\mathbf{P}_e^f$	Ensemble estimate of forecast error covariance matrix
$\mathbf{P}^a$	Analysis error covariance matrix
$\mathbf{P}_e^a$	Ensemble estimate of analysis error covariance matrix
$\mathbf{R}_e$	Observation error covariance matrix
$\boldsymbol{h}$	Non-linear observation operator
$\mathbf{H}$	Tangent linear model of observation operator
$\mathbf{K}$	Kalman gain matrix
$\mathbf{A}^f$	Forecast state vector matrix (forecasts are columns)
$\mathbf{A}'^f$	Forecast state vector perturbation matrix
$\mathbf{A}^a$	Analysis state vector matrix (analyses are columns)
$\mathbf{A}'^a$	Analysis state vector perturbation matrix
$\mathbf{D}$	Observation vector matrix (observation sets are columns)
Υ	Observation vector perturbation matrix
$\mathbf{I}_{N\times N}$	$N \times N$ identity matrix
$\mathbf{1}_{N\times N}$	$N \times N$ matrix of $1/N$ values
$\mathbf{S}$	Model observation perturbations calculated from forecast perturbations
$\mathbf{C}$	$(N-1)\times$ 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

$$\mathbf{P}^f = \overline{(\psi_i^f - \psi^t)(\psi_i^f - \psi^t)^T}$$

$$\mathbf{P}_e^f = \overline{(\psi_i^f - \overline{\psi^f})(\psi_i^f - \overline{\psi^f})^T}$$

Analysis error covariance matrices

$$\mathbf{P}^a = \overline{(\psi_i^a - \psi^t)(\psi_i^a - \psi^t)^T}$$

$$\mathbf{P}_e^a = \overline{(\psi_i^a - \overline{\psi^a})(\psi_i^a - \overline{\psi^a})^T}$$

Observation error covariance matrix

$$\mathbf{R}_e = \overline{(\boldsymbol{y}_i - \overline{\boldsymbol{y}_i})(\boldsymbol{y}_i - \overline{\boldsymbol{y}_i})^T}$$

The observation operator

$$\boldsymbol{y}_i^{(m)} = \boldsymbol{h}(\overline{\psi_i^f}) + \mathbf{H}(\psi_i^f - \overline{\psi_i^f})$$

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

Analysis step for ensemble member *i*

$$\psi_i^a = \psi_i^f + \mathbf{P}_e^f\mathbf{H}^T(\mathbf{H}\mathbf{P}_e^f\mathbf{H}^T + \mathbf{R}_e)^{-1}(\boldsymbol{y}_i - \boldsymbol{h}(\overline{\psi_i^f}) - \mathbf{H}(\psi_i^f - \overline{\psi_i^f}))$$

$$\overline{\psi_i^a} = \overline{\psi_i^f} + \mathbf{P}_e^f\mathbf{H}^T(\mathbf{H}\mathbf{P}_e^f\mathbf{H}^T + \mathbf{R}_e)^{-1}(\bar{\boldsymbol{y}}_i - \boldsymbol{h}(\overline{\psi_i^f}))$$

Deviation of each member from the mean

$$\begin{aligned}\psi_i^a - \overline{\psi_i^a} &= \psi_i^f - \overline{\psi_i^f} + \mathbf{P}_e^f\mathbf{H}^T(\mathbf{H}\mathbf{P}_e^f\mathbf{H}^T + \mathbf{R}_e)^{-1}(\boldsymbol{y}_i - \overline{\boldsymbol{y}_i} - \mathbf{H}(\psi_i^f - \overline{\psi_i^f}))\\ &= (\mathbf{I} - \mathbf{K}\mathbf{H})(\psi_i^f - \overline{\psi_i^f}) + \mathbf{K}(\boldsymbol{y}_i - \overline{\boldsymbol{y}_i})\\ \mathbf{K} &= \mathbf{P}_e^f\mathbf{H}^T(\mathbf{H}\mathbf{P}_e^f\mathbf{H}^T + \mathbf{R}_e)^{-1}\end{aligned}$$

The ensemble analysis error covariance

Assume no correlation between forecast and observation errors.

$$\begin{aligned}\mathbf{P}_e^a &= \left\langle\left[(\mathbf{I} - \mathbf{K}\mathbf{H})(\psi_i^f - \overline{\psi_i^f}) + \mathbf{K}(\boldsymbol{y}_i - \overline{\boldsymbol{y}_i})\right]\left[(\mathbf{I} - \mathbf{K}\mathbf{H})(\psi_i^f - \overline{\psi_i^f}) + \mathbf{K}(\boldsymbol{y}_i - \overline{\boldsymbol{y}_i})\right]^T\right\rangle\\ &= \left\langle\left[(\mathbf{I} - \mathbf{K}\mathbf{H})(\psi_i^f - \overline{\psi_i^f})\right]\left[(\mathbf{I} - \mathbf{K}\mathbf{H})(\psi_i^f - \overline{\psi_i^f})\right]^T\right\rangle + \left\langle\left[\mathbf{K}(\boldsymbol{y}_i - \overline{\boldsymbol{y}_i})\right]\left[\mathbf{K}(\boldsymbol{y}_i - \overline{\boldsymbol{y}_i})\right]^T\right\rangle\\ &= (\mathbf{I} - \mathbf{K}\mathbf{H})\mathbf{P}_e^f(\mathbf{I} - \mathbf{K}\mathbf{H})^T + \mathbf{K}\mathbf{R}_e\mathbf{K}^T\\ &= \mathbf{P}_e^f - \mathbf{P}_e^f\mathbf{H}^T\mathbf{K}^T - \mathbf{K}\mathbf{H}\mathbf{P}_e^f + \mathbf{K}\mathbf{H}\mathbf{P}_e^f\mathbf{H}^T\mathbf{K}^T + \mathbf{K}\mathbf{R}_e\mathbf{K}^T\\ &= \mathbf{P}_e^f + \mathbf{K}(\mathbf{H}\mathbf{P}_e^f\mathbf{H}^T + \mathbf{R}_e)\mathbf{K}^T - \mathbf{P}_e^f\mathbf{H}^T\mathbf{K}^T - \mathbf{K}\mathbf{H}\mathbf{P}_e^f\\ &= \mathbf{P}_e^f + \mathbf{P}_e^f\mathbf{H}^T(\mathbf{H}\mathbf{P}_e^f\mathbf{H}^T + \mathbf{R}_e)^{-1}(\mathbf{H}\mathbf{P}_e^f\mathbf{H}^T + \mathbf{R}_e)\mathbf{K}^T - \mathbf{P}_e^f\mathbf{H}^T\mathbf{K}^T - \mathbf{K}\mathbf{H}\mathbf{P}_e^f\end{aligned}$$

$$= \mathbf{P}_e^f - \mathbf{KHP}_e^f$$

$$= (\mathbf{I} - \mathbf{KH})\,\mathbf{P}_e^f$$

Matrix representation of the ensemble

The forecast states ψ_i^f may be assembled into columns of the $n \times N$ matrix $\mathbf{A}^f$.

$$\mathbf{A}^f = (\psi_1^f\,\psi_2^f \ldots\ \psi_N^f)$$

The analysis states ψ_i^a may be assembled into columns of the $n \times N$ matrix $\mathbf{A}^a$.

$$\mathbf{A}^a = (\psi_1^a\,\psi_2^a \ldots\ \psi_N^a)$$

The perturbed observations $\boldsymbol{y}_i$ may be assembled into the columns of the $p \times N$ matrix $\mathbf{D}$.

$$\mathbf{D} = (\boldsymbol{y}_1\,\boldsymbol{y}_2 \ldots\ \boldsymbol{y}_N)$$

The ensemble means, $\bar{\mathbf{A}}^f$, $\bar{\mathbf{A}}^a$ and $\bar{\mathbf{D}}$ are matrices that comprise the respective ensemble mean state repeated in each column.

$$\bar{\mathbf{A}}^f = \mathbf{A}^f\mathbf{1}_{N\times N} \qquad \bar{\mathbf{A}}^a = \mathbf{A}^a\mathbf{1}_{N\times N} \qquad \bar{\mathbf{D}} = \mathbf{D1}_{N\times N}$$

$$\mathbf{1}_{N\times N} = \begin{pmatrix} 1/N & 1/N & \ldots & 1/N \\ 1/N & 1/N & \ldots & 1/N \\ \ldots & \ldots & \ldots & \ldots \\ 1/N & 1/N & \ldots & 1/N \end{pmatrix}$$

Matrix representation of the ensemble perturbations from the mean

$$\mathbf{A}'^f = \mathbf{A}^f - \bar{\mathbf{A}}^f = \mathbf{A}^f(\mathbf{I}_{N\times N} - \mathbf{1}_{N\times N})$$

$$\mathbf{A}'^a = \mathbf{A}^a - \bar{\mathbf{A}}^a = \mathbf{A}^a(\mathbf{I}_{N\times N} - \mathbf{1}_{N\times N})$$

$$\Upsilon = \mathbf{D} - \bar{\mathbf{D}} = \mathbf{D}(\mathbf{I}_{N\times N} - \mathbf{1}_{N\times N})$$

Note: $\quad (\mathbf{I}_{N\times N} - \mathbf{1}_{N\times N})^2 = (\mathbf{I}_{N\times N} - \mathbf{1}_{N\times N})$

The error covariance matrices from the matrix representation of ensemble perturbations

$$\mathbf{P}_e^f = \frac{1}{N-1}\mathbf{A}'^f\mathbf{A}'^{fT} \qquad \mathbf{P}_e^a = \frac{1}{N-1}\mathbf{A}'^a\mathbf{A}'^{aT} \qquad \mathbf{R}_e = \frac{1}{N-1}\Upsilon\Upsilon^T$$

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

$$\mathbf{A}^a = \mathbf{A}^f + \mathbf{K}(\mathbf{D} - \mathbf{HA}^f)$$

$$\mathbf{K} = \mathbf{P}_e^f\mathbf{H}^T(\mathbf{HP}_e^f\mathbf{H}^T + \mathbf{R}_e)^{-1}$$

$$= \mathbf{A}'^f\mathbf{A}'^{fT}\mathbf{H}^T(\mathbf{HA}'^f\mathbf{A}'^{fT}\mathbf{H}^T + \Upsilon\Upsilon^T)^{-1}$$

$$= \mathbf{A}'^f(\mathbf{HA}'^f)^T((\mathbf{HA}'^f)(\mathbf{HA}'^f)^T + \Upsilon\Upsilon^T)^{-1}$$

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

$$\bar{\mathbf{A}}^a = \mathbf{A}^a \mathbf{1}_{N\times N} = \mathbf{A}^f \mathbf{1}_{N\times N} + \mathbf{K}(\mathbf{D} - \mathbf{H}\mathbf{A}^f)\mathbf{1}_{N\times N}$$
$$= \bar{\mathbf{A}}^f + \mathbf{K}(\bar{\mathbf{D}} - \mathbf{H}\bar{\mathbf{A}}^f)$$

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

$$\mathbf{P}_e^a = \frac{1}{N-1}\mathbf{A}'^a\mathbf{A}'^{a^T} = \frac{1}{N-1}\mathbf{A}^a(\mathbf{I}_{N\times N} - \mathbf{1}_{N\times N})(\mathbf{I}_{N\times N} - \mathbf{1}_{N\times N})\mathbf{A}^{a^T}$$
$$= \frac{1}{N-1}\mathbf{A}^a(\mathbf{I}_{N\times N} - \mathbf{1}_{N\times N})\mathbf{A}^{a^T}$$
$$= \frac{1}{N-1}\left[\mathbf{A}^f + \mathbf{K}(\mathbf{D} - \mathbf{H}\mathbf{A}^f)\right](\mathbf{I}_{N\times N} - \mathbf{1}_{N\times N})\left[\mathbf{A}^f + \mathbf{K}(\mathbf{D} - \mathbf{H}\mathbf{A}^f)\right]^T$$
$$= \mathbf{P}_e^f - \mathbf{P}_e^f\mathbf{H}^T\mathbf{K}^T - \mathbf{K}\mathbf{H}\mathbf{P}_e^f + \mathbf{K}\mathbf{R}_e\mathbf{K}^T + \mathbf{K}\mathbf{H}\mathbf{P}_e^f\mathbf{H}^T\mathbf{K}^T$$
$$= (\mathbf{I} - \mathbf{K}\mathbf{H})\mathbf{P}_e^f \qquad \text{using analysis beforehand.}$$

Factorization in terms of S and C matrices

The ensemble mean analysis can be written in terms of matrices **S** and **C**.

$$\bar{\mathbf{A}}^a = \bar{\mathbf{A}}^f + \mathbf{A}'^f\mathbf{S}^T\mathbf{C}^{-1}(\bar{\mathbf{D}} - \mathbf{H}\bar{\mathbf{A}}^f)$$
$$\text{where} \qquad \mathbf{S} = \mathbf{H}\mathbf{A}'^f \qquad \mathbf{C} = \mathbf{S}\mathbf{S}^T + (N-1)\mathbf{R}_e$$

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, $\mathbf{S}\mathbf{S}^T$, and observation error covariances, $\mathbf{R}_e$.

The above result can be checked by substitution

$$\bar{\mathbf{A}}^a = \bar{\mathbf{A}}^f + \mathbf{A}'^f\mathbf{A}'^{f^T}\mathbf{H}^T\left[\mathbf{H}\mathbf{A}'^f\mathbf{A}'^{f^T}\mathbf{H}^T + (N-1)\mathbf{R}_e\right]^{-1}(\bar{\mathbf{D}} - \mathbf{H}\bar{\mathbf{A}}^f)$$
$$= \bar{\mathbf{A}}^f + (N-1)\mathbf{P}_e^f\mathbf{H}^T\left[\mathbf{H}(N-1)\mathbf{P}_e^f\mathbf{H}^T + (N-1)\mathbf{R}_e\right]^{-1}(\bar{\mathbf{D}} - \mathbf{H}\bar{\mathbf{A}}^f)$$
$$= \bar{\mathbf{A}}^f + \mathbf{P}_e^f\mathbf{H}^T\left[\mathbf{H}\mathbf{P}_e^f\mathbf{H}^T + \mathbf{R}_e\right]^{-1}(\bar{\mathbf{D}} - \mathbf{H}\bar{\mathbf{A}}^f)$$
$$= \bar{\mathbf{A}}^f + \mathbf{K}(\bar{\mathbf{D}} - \mathbf{H}\bar{\mathbf{A}}^f) \qquad \text{expression is confirmed.}$$

$\mathbf{P}_e^a$ can also be written in terms of matrices **S** and **C**.

$$\mathbf{P}_e^a = \frac{1}{N-1}\mathbf{A}'^f(\mathbf{I} - \mathbf{S}^T\mathbf{C}^{-1}\mathbf{S})\mathbf{A}'^{f^T}$$

This result can also be checked by substitution

$$\mathbf{P}_e^a = \frac{1}{N-1}\mathbf{A}'^f\mathbf{A}'^{f^T} - \frac{1}{N-1}\mathbf{A}'^f\mathbf{A}'^{f^T}\mathbf{H}^T\left[\mathbf{H}\mathbf{A}'^f\mathbf{A}'^{f^T}\mathbf{H}^T + (N-1)\mathbf{R}_e\right]^{-1}\mathbf{H}\mathbf{A}'^f\mathbf{A}'^{f^T}$$
$$= \mathbf{P}_e^f - \mathbf{P}_e^f\mathbf{H}^T\left[\mathbf{H}(N-1)\mathbf{P}_e^f\mathbf{H}^T + (N-1)\mathbf{R}_e\right]^{-1}(N-1)\mathbf{H}\mathbf{P}_e^f$$
$$= \mathbf{P}_e^f - \mathbf{P}_e^f\mathbf{H}^T\left[\mathbf{H}\mathbf{P}_e^f\mathbf{H}^T + \mathbf{R}_e\right]^{-1}\mathbf{H}\mathbf{P}_e^f$$

$$= \mathbf{P}_e^f - \mathbf{K}\mathbf{H}\mathbf{P}_e^f$$

$$= (\mathbf{I} - \mathbf{K}\mathbf{H})\,\mathbf{P}_e^f \qquad \text{expression is confirmed.}$$

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 $\mathbf{P}_e^a$. The square-root matrix may be regarded as a set of ensemble perturbations that have covariance $\mathbf{P}_e^a$. Given that **C** is available, write it in terms of its eigenvalue decomposition

$$\mathbf{C} = \mathbf{Z}\Lambda\mathbf{Z}^T$$

where Λ 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. $\mathbf{P}_e^a$ is then written

$$\mathbf{P}_e^a = \frac{1}{N-1}\mathbf{A}'^f\,(\mathbf{I} - \mathbf{S}^T\mathbf{Z}\Lambda^{-1}\mathbf{Z}^T\mathbf{S})\,\mathbf{A}'^{f^T}$$

$$= \frac{1}{N-1}\mathbf{A}'^f\,(\mathbf{I} - \mathbf{X}^T\mathbf{X})\,\mathbf{A}'^{f^T}$$

$$\text{where} \qquad \mathbf{X} = \Lambda^{-1/2}\mathbf{Z}^T\mathbf{S}$$

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

$$\mathbf{X} = \mathbf{U}\Sigma\mathbf{V}^T$$

where Σ 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. $\mathbf{P}_e^a$ is then written

$$\mathbf{P}_e^a = \frac{1}{N-1}\mathbf{A}'^f\,(\mathbf{I} - \mathbf{V}\Sigma^T\mathbf{U}^T\mathbf{U}\Sigma\mathbf{V}^T)\,\mathbf{A}'^{f^T}$$

$$= \frac{1}{N-1}\mathbf{A}'^f\,(\mathbf{I} - \mathbf{V}\Sigma^T\Sigma\mathbf{V}^T)\,\mathbf{A}'^{f^T}$$

$$= \frac{1}{N-1}\mathbf{A}'^f\mathbf{V}\,(\mathbf{I} - \Sigma^T\Sigma)\,\mathbf{V}^T\mathbf{A}'^{f^T}$$

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

$$\mathbf{P}_e^a = \frac{1}{N-1}\mathbf{A}'^a\mathbf{A}'^{a^T} = \frac{1}{N-1}\mathbf{A}'^f\mathbf{V}\,(\mathbf{I} - \Sigma^T\Sigma)^{1/2}\,(\mathbf{I} - \Sigma^T\Sigma)^{1/2}\,\mathbf{V}^T\mathbf{A}'^{f^T}$$

$$\therefore \qquad \mathbf{A}'^a\mathbf{A}'^{a^T} = \left[\mathbf{A}'^f\mathbf{V}\,(\mathbf{I} - \Sigma^T\Sigma)^{1/2}\right]\left[\mathbf{A}'^f\mathbf{V}\,(\mathbf{I} - \Sigma^T\Sigma)^{1/2}\right]^T \qquad \text{and}$$

$$\mathbf{A}'^a = \mathbf{A}'^f\mathbf{V}\,(\mathbf{I} - \Sigma^T\Sigma)^{1/2}$$