

A Reduced Rank Kalman Filter

Ross Bannister, October/November 2009

These notes give the detailed workings of Fisher's reduced rank Kalman filter (RRKF), which was developed for use in a variational environment (Fisher, 1998). These notes are my interpretation of the method and an outline of how I would implement it. Fisher's approach uses flow-dependent information from a Hessian singular vector (HSV) calculation and I propose an alternative that uses information from an available ensemble, made up of ensemble members (EMs).

1. What is a reduced rank Kalman filter?

In VAR, the **B**-matrix used is a static representation of forecast error covariances and as such VAR would be expected to perform poorly in an environment where the actual forecast error covariance statistics change significantly from case to case (e.g. Fisher, 2007; Bannister, 2008). The RRKF is an attempt to blend-in to the Var. problem flow-dependent aspects in a mathematically formal manner. Two sources of flow-dependent information are considered here: firstly HSV information (as in Fisher's original formulation (Fisher, 1998)) and secondly from EM information.

A RRKF in this context may be regarded as a modification to the existing **B**-matrix in VAR that allows the dynamical evolution of a subspace of the state vector (where the subspace is defined as that spanned by either the HSVs or by the EMs). Each approach is outlined below.

2. Source A of flow-dependent information to blend with the B-matrix (Hessian singular vectors)

One way of identifying the subspace that will be treated with explicit flow dependence is to use a HSV calculation. Let the size of the subspace be K , which can be chosen arbitrarily, but restricted in practice by cost. K represents the number of singular vectors. The size of the full model space is N and in these notes it is recognised that $K \ll N$.

Fisher defines the subspace by the K most rapidly growing HSVs. The reason why they are chosen to be singular vectors of the Hessian will become evident. In order to specify the problem that must be solved to compute the HSVs, we introduce two norms as follows.

- Let the covariance matrix \mathbf{P}^a be the error covariance of the analysis of the previous cycle. In order for Fisher's method to work, it must be possible to act with the matrix \mathbf{P}^{a-1} (or an approximation of it).
- Let the matrix \mathbf{W} be the norm used to measure the size of a perturbation. It must be possible to act with the matrix \mathbf{W}^{-1} .

Let the time of the previous data assimilation cycle be $-t$ and the time of the present analysis be 0. States that have no time label are valid at time 0 by default.

Let the tangent linear model, $\mathbf{M}_{0 \leftarrow -t}$ act on perturbations at time $-t$ and give a perturbation at time 0

$$\delta\mathbf{x} = \mathbf{M}_{0 \leftarrow -t}\delta\mathbf{x}(-t). \quad (1)$$

If $\delta\mathbf{x}(-t)$ were known, then the size of $\delta\mathbf{x}$ according to the \mathbf{W} -norm would be J_1

$$J_1 = \delta\mathbf{x}^T \mathbf{W}^{-1} \delta\mathbf{x} = \delta\mathbf{x}^T(-t) \mathbf{M}_{0 \leftarrow -t}^T \mathbf{W}^{-1} \mathbf{M}_{0 \leftarrow -t} \delta\mathbf{x}(-t). \quad (2)$$

The HSVs are defined as those $\delta\mathbf{x}(-t)$ that maximise J_1 subject to the constraint that $\delta\mathbf{x}(-t)$ is distributed according to \mathbf{P}^a , ie

$$\delta\mathbf{x}^T(-t) \mathbf{P}^{a-1} \delta\mathbf{x}(-t) - \text{const} = 0, \quad (3)$$

for an arbitrary constant, 'const'. The constrained optimisation problem may therefore be posed as

$$\nabla_{\mathbf{x}(-t)} [J_1 - \lambda (\delta\mathbf{x}^T(-t) \mathbf{P}^{a-1} \delta\mathbf{x}(-t) - \text{const})] = 0, \quad (4)$$

where λ is the Lagrange multiplier. Differentiating and setting the solutions to $\delta\mathbf{x}_k(-t)$ (with associated Lagrange multiplier λ_k) gives

$$\mathbf{M}_{0 \leftarrow -t}^T \mathbf{W}^{-1} \mathbf{M}_{0 \leftarrow -t} \delta\mathbf{x}_k(-t) = \lambda_k \mathbf{P}^{a-1} \delta\mathbf{x}_k(-t), \quad (5)$$

which is a generalised eigenvalue problem. The $\delta\mathbf{x}_k(-t)$ are the HSVs. The set of vectors $\mathbf{P}^{a-1/2} \delta\mathbf{x}_k(-t)$ are eigenvectors of $(\mathbf{W}^{-1/2} \mathbf{M}_{0 \leftarrow -t} \mathbf{P}^{a/2})^T (\mathbf{W}^{-1/2} \mathbf{M}_{0 \leftarrow -t} \mathbf{P}^{a/2})$ and are thus the right singular vectors of the matrix $\mathbf{W}^{-1/2} \mathbf{M}_{0 \leftarrow -t} \mathbf{P}^{a/2}$. Let $\mathbf{s}_k = \mathbf{M}_{0 \leftarrow -t} \delta\mathbf{x}_k(-t)$. Those \mathbf{s}_k with the largest λ_k define the subspace whose background errors are treated explicitly by the RRKF.

A general perturbation at time 0, $\delta\mathbf{x}$ has a part $\delta\mathbf{x}_s$ that lies in this subspace, which can be found as a linear combination of the \mathbf{s}_k . Identification of this subspace can be simplified by first constructing an orthogonalized and normalized set of vectors, $\tilde{\mathbf{s}}_k$ (e.g. by the Gramm-Schmidt procedure, see App. A). Then

$$\delta\mathbf{x}_s = \tilde{\mathbf{S}} \mathbf{a}, \quad (6)$$

where $\tilde{\mathbf{S}}$ is the $N \times K$ matrix of $\tilde{\mathbf{s}}_k$ vectors and \mathbf{a} is the K -element vector of (as yet unknown) coefficients. Orthogonalization should be done with respect to an inner product that non-dimensionalizes the components (the \mathbf{W}^{-1} inner product achieves this, so this matrix shall be used) such that

$$\tilde{\mathbf{S}}^T \mathbf{W}^{-1} \tilde{\mathbf{S}} = \mathbf{I}, \quad (7)$$

(the \mathbf{B}^{-1} norm could, in principle be used instead). The benefit of first orthogonalizing the vectors is to allow \mathbf{a} to be found easily from $\delta\mathbf{x}$

$$\mathbf{a} = \tilde{\mathbf{S}}^T \mathbf{W}^{-1} \delta\mathbf{x}. \quad (8)$$

The part of $\delta\mathbf{x}$ that is not within the chosen subspace is the residual $\delta\bar{\mathbf{x}}$,

$$\delta\bar{\mathbf{x}}_s = \delta\mathbf{x} - \delta\mathbf{x}_s, \quad (9)$$

which is orthogonal to $\delta\mathbf{x}_s$ under the \mathbf{W}^{-1} norm (see App. B).

The way that the HSVs may be used in the RRKF is covered in Sec. 6.

3. Source B of flow-dependent information to blend with the B-matrix (ensemble members)

Another way of identifying the subspace that will be treated with explicit flow dependence is to use information extracted from an ensemble. Let the number of EMs be L , which can be chosen arbitrarily, but restricted in practice by cost. From the L ensemble members, a K -dimensional subspace can be determined ($K \leq L$) that will be used to describe the covariances explicitly.

The size of the full model space is N and in these notes it is assumed that $K, L \ll N$.

In a similar way to the procedure for the HSVs in Sec. 2, the EMs may be orthogonalized by the Gram-Schmidt procedure (see App. A) and placed in columns of the matrix $\tilde{\mathbf{S}}$. The relationship between a vector in the L -dimensional space spanned by the EMs, \mathbf{a} , and the model space, $\delta\mathbf{x}_s$, is given by (6), but where now \mathbf{a} is a L -element vector and $\tilde{\mathbf{S}}$ is the $N \times L$ matrix of orthogonalized EMs. Orthogonalization is performed with respect to the \mathbf{W}^{-1} inner product described by (7) (Eqs. (8) and (9) then follow).

The procedure to use the EMs in the RRKF is a little more involved than the HSV procedure. The way that the EMs may be used in the RRKF is covered in Sec. 7.

4. The background cost function in the new subspace

The next task is to outline the way that the flow-dependent information, whether from HSVs or from EMs, can be combined with static error covariance information in the VAR formulation. The usual background cost function in VAR is J_b

$$J_b = \frac{1}{2} (\delta\mathbf{x} - \delta\mathbf{x}^b)^T \mathbf{B}^{-1} (\delta\mathbf{x} - \delta\mathbf{x}^b), \quad (10)$$

where $\delta\mathbf{x} = \mathbf{x} - \mathbf{x}^g$, $\delta\mathbf{x}^b = \mathbf{x}^b - \mathbf{x}^g$, \mathbf{x}^b is the background state and \mathbf{x}^g is a reference (or guess) state. The \mathbf{B} -matrix in (10) is static. Equation (10) may be written in terms of the components $\delta\mathbf{x}_s$ and $\delta\bar{\mathbf{x}}_s$ by substituting (9) into (10). This gives three parts: (i) a part that involves only the special subspace that has been identified from the K -dimensional subspace introduced in Secs. 2 and 3, (ii) the part that couples this subspace with the rest of the state, and (iii) the part that involves only the rest of the state

$$J_b = \frac{1}{2} (\delta\mathbf{x}_s - \delta\mathbf{x}_s^b)^T \mathbf{B}^{-1} (\delta\mathbf{x}_s - \delta\mathbf{x}_s^b) + (\delta\bar{\mathbf{x}}_s - \delta\bar{\mathbf{x}}_s^b)^T \mathbf{B}^{-1} (\delta\mathbf{x}_s - \delta\mathbf{x}_s^b) + \frac{1}{2} (\delta\bar{\mathbf{x}}_s - \delta\bar{\mathbf{x}}_s^b)^T \mathbf{B}^{-1} (\delta\bar{\mathbf{x}}_s - \delta\bar{\mathbf{x}}_s^b). \quad (11)$$

This cost function is identical to (10). The RRKF is constructed by imposing a flow dependent error covariance matrix for the first two terms ($\mathbf{B} \rightarrow \mathbf{P}^f$) but keeping the static \mathbf{B} -matrix in the last term

$$J_b \rightarrow \frac{1}{2} (\delta\mathbf{x}_s - \delta\mathbf{x}_s^b)^T \mathbf{P}^{f-1} (\delta\mathbf{x}_s - \delta\mathbf{x}_s^b) + \alpha (\delta\bar{\mathbf{x}}_s - \delta\bar{\mathbf{x}}_s^b)^T \mathbf{P}^{f-1} (\delta\mathbf{x}_s - \delta\mathbf{x}_s^b) + \frac{1}{2} (\delta\bar{\mathbf{x}}_s - \delta\bar{\mathbf{x}}_s^b)^T \mathbf{B}^{-1} (\delta\bar{\mathbf{x}}_s - \delta\bar{\mathbf{x}}_s^b). \quad (12)$$

The factor α , added by Fisher (1998), is to help ensure that J_b is convex. The flow dependent information provided by the K special vectors will be used to determine the \mathbf{P}^f . It will be introduced into the problem by a modification to the standard control variable transform used in VAR.

5. Control variable transforms stage

It is usual in VAR to make a change of variable from model variables ($\delta\mathbf{x}$) to control variables (often named χ). The control variable transform used in standard VAR is here denoted \mathbf{L} and in the RRKF there is an additional transform denoted \mathbf{X} as follows

$$\delta \mathbf{x} = \mathbf{LX}\boldsymbol{\chi}, \quad (13)$$

where, by design, \mathbf{X} is an orthogonal matrix

$$\mathbf{XX}^T = \mathbf{I}, \quad (14)$$

(see below) and \mathbf{L} is the usual control variable transform used in VAR

$$\mathbf{LX}(\mathbf{LX})^T = \mathbf{LL}^T = \mathbf{B}. \quad (15)$$

Substituting (13) into (12) gives

$$\begin{aligned} J_b = & \frac{1}{2}(\boldsymbol{\chi}_s - \boldsymbol{\chi}_s^b)^T \mathbf{X}^T \mathbf{L}^T \mathbf{P}^{f-1} \mathbf{LX} (\boldsymbol{\chi}_s - \boldsymbol{\chi}_s^b) + \alpha (\bar{\boldsymbol{\chi}}_s - \bar{\boldsymbol{\chi}}_s^b)^T \mathbf{X}^T \mathbf{L}^T \mathbf{P}^{f-1} \mathbf{LX} (\boldsymbol{\chi}_s - \boldsymbol{\chi}_s^b) + \\ & \frac{1}{2}(\bar{\boldsymbol{\chi}}_s - \bar{\boldsymbol{\chi}}_s^b)^T \mathbf{X}^T \mathbf{L}^T \mathbf{B}^{-1} \mathbf{LX} (\bar{\boldsymbol{\chi}}_s - \bar{\boldsymbol{\chi}}_s^b), \end{aligned} \quad (16)$$

where $\boldsymbol{\chi}_s = \mathbf{X}^T \mathbf{L}^{-1} \delta \mathbf{x}_s$, $\boldsymbol{\chi}_s^b = \mathbf{X}^T \mathbf{L}^{-1} \delta \mathbf{x}_s^b$, $\bar{\boldsymbol{\chi}}_s = \mathbf{X}^T \mathbf{L}^{-1} \delta \bar{\mathbf{x}}_s$ and $\bar{\boldsymbol{\chi}}_s^b = \mathbf{X}^T \mathbf{L}^{-1} \delta \bar{\mathbf{x}}_s^b$. The matrix \mathbf{X} is not present in standard VAR, but is introduced in (13) to isolate the special subspace identified in Secs. 2 and 3 from the remainder of the state space. As it stands, (16) looks complicated to treat. Two substitutions are made as follows

$$\mathbf{X}^T \mathbf{L}^T \mathbf{B}^{-1} \mathbf{LX} = \mathbf{I}, \quad (17)$$

$$\mathbf{X}^T \mathbf{L}^T \mathbf{P}^{f-1} \mathbf{LX} = \mathbf{P}_\chi^{f-1}, \quad (18)$$

where (17) follows from (14) and (15), and (18) is a definition. With these substitutions, (16) is

$$\begin{aligned} J_b = & \frac{1}{2}(\boldsymbol{\chi}_s - \boldsymbol{\chi}_s^b)^T \mathbf{P}_\chi^{f-1} (\boldsymbol{\chi}_s - \boldsymbol{\chi}_s^b) + \alpha (\bar{\boldsymbol{\chi}}_s - \bar{\boldsymbol{\chi}}_s^b)^T \mathbf{P}_\chi^{f-1} (\boldsymbol{\chi}_s - \boldsymbol{\chi}_s^b) + \\ & \frac{1}{2}(\bar{\boldsymbol{\chi}}_s - \bar{\boldsymbol{\chi}}_s^b)^T (\bar{\boldsymbol{\chi}}_s - \bar{\boldsymbol{\chi}}_s^b). \end{aligned} \quad (19)$$

The part of \mathbf{P}_χ^{f-1} that is important is derived in Secs. 6 and 7 from the K -dimensional subspaces identified from the HSV or EM calculations respectively. The key to simplifying (19) is in the design of \mathbf{X} . Let \mathbf{X} have the following properties.

- \mathbf{X}^T acting on any vector in the subspace $\mathbf{L}^{-1} \delta \mathbf{x}_s$ (where $\delta \mathbf{x}_s$ is a vector that exists entirely in the special K -dimensional subspace) gives a vector that is non-zero only in the first K elements

$$\mathbf{X}^T \mathbf{L}^{-1} \delta \mathbf{x}_s = \begin{pmatrix} \alpha_1 \\ \dots \\ \alpha_K \\ 0 \\ \dots \\ 0 \end{pmatrix}, \quad (20)$$

- and \mathbf{X}^T acting on a vector in the remaining space, $\mathbf{L}^{-1} \delta \bar{\mathbf{x}}_s$, gives a vector that is non-zero only in the remaining $N - K$ elements.

It is possible to define a suitable \mathbf{X} that satisfies these conditions by using a sequence of Householder transformations (see App. C). The important observation is that in (19), only the first K columns of \mathbf{P}_χ^{f-1} need to be known.

6. Determination of \mathbf{P}_χ^{f-1} for the Hessian singular vectors

The procedure to calculate the first K columns of \mathbf{P}_χ^{f-1} using the HSVs is now described.

Following Fisher (1998) let

$$\mathbf{Z} = \mathbf{P}^{f^{-1}} \mathbf{S}, \quad (21)$$

where \mathbf{S} is the $N \times K$ matrix whose columns are the s_k and \mathbf{Z} is the $N \times K$ result after acting with the inverse of the flow-dependent error covariance matrix. Equation (21) is now developed using the definition (18) along the way

$$\begin{aligned} \mathbf{Z} &= \mathbf{P}^{f^{-1}} \mathbf{L} \mathbf{X} \mathbf{X}^T \mathbf{L}^{-1} \mathbf{S}, \\ \mathbf{X}^T \mathbf{L}^T \mathbf{Z} &= \mathbf{X}^T \mathbf{L}^T \mathbf{P}^{f^{-1}} \mathbf{L} \mathbf{X} \mathbf{X}^T \mathbf{L}^{-1} \mathbf{S}, \\ &= \mathbf{P}_{\chi}^{f^{-1}} \hat{\mathbf{I}} \mathbf{X}^T \mathbf{L}^{-1} \mathbf{S}, \end{aligned} \quad (22)$$

$$\mathbf{X}_{(N \times N)}^T \mathbf{L}_{(N \times N)}^T \mathbf{Z}_{(N \times K)} = \mathbf{P}_{\chi(N \times K)}^{f^{-1}} \hat{\mathbf{I}}_{(K \times N)} \mathbf{X}_{(N \times N)}^T \mathbf{L}_{(N \times N)}^{-1} \mathbf{S}_{(N \times K)}, \quad (23)$$

where $\hat{\mathbf{I}}$ is the following non-square quasi-identity matrix

$$\hat{\mathbf{I}}_{(K \times N)} = (\mathbf{I}_{(K \times K)} \mathbf{0}_{(K \times N - K)}). \quad (24)$$

This matrix is included to remove the superfluous zero-elements for rows $i > K$ of $\mathbf{X}^T \mathbf{L}^{-1} \mathbf{S}$ (by the design of \mathbf{X}). In (23) and (24), labels have been added to the matrices to indicate their dimensions. Equation (22) leads to

$$\mathbf{X}^T \mathbf{L}^T \mathbf{Z} (\hat{\mathbf{I}} \mathbf{X}^T \mathbf{L}^{-1} \mathbf{S})^{-1} = \mathbf{P}_{\chi}^{f^{-1}}, \quad (25)$$

where the operator inverted is a calculable $K \times K$ matrix, which we assume is non-singular. Note that (25) is for only part of the inverse covariance matrix and so is not symmetric. The matrix yet known is $\mathbf{X}^T \mathbf{L}^T \mathbf{Z}$ which is now found from the HSVs (Sec. 2).

By the definition of \mathbf{Z} (21), $\mathbf{X}^T \mathbf{L}^T \mathbf{Z}$ is

$$\mathbf{X}^T \mathbf{L}^T \mathbf{Z} = \mathbf{X}^T \mathbf{L}^T \mathbf{P}^{f^{-1}} \mathbf{S}, \quad (26)$$

whose right hand side can be found from (5). Let columns of a new matrix, \mathbf{S}_{-t} , be those K vectors at time t that evolve into the columns of \mathbf{S} (the columns of \mathbf{S}_{-t} are the states $\delta \mathbf{x}_k(-t)$ in (5))

$$\mathbf{S} = \mathbf{M}_{0 \leftarrow -t} \mathbf{S}_{-t}. \quad (27)$$

This is useful in the derivation to follow. Write (5) in complete matrix form

$$\mathbf{M}_{0 \leftarrow -t}^T \mathbf{W}^{-1} \mathbf{M}_{0 \leftarrow -t} \mathbf{S}_{-t} = \mathbf{P}^a \mathbf{S}_{-t} \Lambda, \quad (28)$$

where Λ is the diagonal matrix of λ_k . Also important is the propagation of the error covariances (ignoring the model error contribution)

$$\mathbf{P}^f = \mathbf{M}_{0 \leftarrow -t} \mathbf{P}^a \mathbf{M}_{0 \leftarrow -t}^T. \quad (29)$$

These equations can be manipulated to give the matrix in (26) required to complete (25). Starting from (28)

$$\begin{aligned} \mathbf{P}^a \mathbf{M}_{0 \leftarrow -t}^T \mathbf{W}^{-1} \mathbf{S} \Lambda^{-1} &= \mathbf{S}_{-t}, \\ \mathbf{M}_{0 \leftarrow -t} \mathbf{P}^a \mathbf{M}_{0 \leftarrow -t}^T \mathbf{W}^{-1} \mathbf{S} \Lambda^{-1} &= \mathbf{M}_{0 \leftarrow -t} \mathbf{S}_{-t}, \\ \mathbf{P}^f \mathbf{W}^{-1} \mathbf{S} \Lambda^{-1} &= \mathbf{S}, \\ \mathbf{W}^{-1} \mathbf{S} \Lambda^{-1} &= \mathbf{P}^{f^{-1}} \mathbf{S}, \\ \mathbf{X}^T \mathbf{L}^T \mathbf{W}^{-1} \mathbf{S} \Lambda^{-1} &= \mathbf{X}^T \mathbf{L}^T \mathbf{P}^{f^{-1}} \mathbf{S}, \end{aligned}$$

$$= \mathbf{X}^T \mathbf{L}^T \mathbf{Z} \quad \text{by (26),}$$

$$\therefore \mathbf{P}_\chi^{f^{-1}} = \mathbf{X}^T \mathbf{L}^T \mathbf{W}^{-1} \mathbf{S} \Lambda^{-1} (\hat{\mathbf{I}} \mathbf{X}^T \mathbf{L}^{-1} \mathbf{S})^{-1} \quad \text{by (25).} \quad (30)$$

The right hand side of (30) is known and thus all relevant elements of the background cost function (19) are now calculable given the HSV results.

7. Completing the calculation with information from the ensemble

The EMs are now used to determine the first K columns of $\mathbf{P}_\chi^{f^{-1}}$. Let the columns of \mathbf{S} contain the L raw EMs ($L \geq K$). The forecast error covariance matrix in state space is then

$$\mathbf{P}^f = \frac{1}{L-1} \mathbf{S} \mathbf{S}^T, \quad (31)$$

which is too large to compute explicitly. The transform between the ensemble from the orthogonalized ensemble subspace and the state space is

$$\mathbf{S} = \tilde{\mathbf{S}} \mathbf{S}_{\text{sub}}, \quad (32)$$

where $\tilde{\mathbf{S}}$ is the $N \times L$ orthogonalized matrix of EMs, as used in (6), and \mathbf{S}_{sub} is the $L \times L$ matrix of EMs in the orthogonalized ensemble representation. The orthogonality of $\tilde{\mathbf{S}}$ is specified in (7). The inverse of (32) is

$$\mathbf{S}_{\text{sub}} = \tilde{\mathbf{S}}^T \mathbf{W}^{-1} \mathbf{S}. \quad (33)$$

The ensemble forecast error covariance matrix in EM subspace is

$$\mathbf{P}_L^f = \frac{1}{L-1} \mathbf{S}_{\text{sub}} \mathbf{S}_{\text{sub}}^T, \quad (34)$$

which is easily calculable, as are its eigenvectors and eigenvalues. For reference, the relationship between \mathbf{P}^f in (31) and \mathbf{P}_L^f in (34) is, using (32)

$$\mathbf{P}^f = \tilde{\mathbf{S}} \mathbf{P}_L^f \tilde{\mathbf{S}}^T. \quad (35)$$

Interest here is in the K eigenvectors of \mathbf{P}_L^f with the largest eigenvalues, where $0 < K \leq L$. These are used to define the special K -dimensional subspace. First let \mathbf{U}_L be the $L \times L$ matrix of all L eigenvectors, where

$$\mathbf{U}_L^T \mathbf{U}_L = \mathbf{I} \quad \text{and} \quad \mathbf{U}_L \mathbf{U}_L^T = \mathbf{I}, \quad (36)$$

and let Λ_L be the diagonal $L \times L$ matrix of eigenvalues. Equation (34) may be decomposed as

$$\mathbf{P}_L^f = \mathbf{U}_L \Lambda_L \mathbf{U}_L^T. \quad (37)$$

Similarly, let \mathbf{U}_K be the $L \times K$ matrix of the K eigenvectors with the largest eigenvalues, where

$$\mathbf{U}_K^T \mathbf{U}_K = \mathbf{I}. \quad (38)$$

The following also holds as long as it acts only on vectors entirely in the column space of \mathbf{U}_K

$$\mathbf{U}_K \mathbf{U}_K^T = \mathbf{I}. \quad (39)$$

The K largest eigenvalues are assembled into the diagonal matrix Λ_K . The $L \times L$ covariance matrix, but of rank K is then (by analogy to (37))

$$\mathbf{P}_K^f = \mathbf{U}_K \Lambda_K \mathbf{U}_K^T. \quad (40)$$

These K eigenvectors may be projected back to model space using (6)

$$\mathbf{U} = \tilde{\mathbf{S}} \mathbf{U}_K, \quad (41)$$

where (as explained in Sec. 3) $\tilde{\mathbf{S}}$ is the $N \times L$ matrix of orthogonalized ensemble members. The inverse of (41) makes use of (7)

$$\mathbf{U}_K = \tilde{\mathbf{S}}^T \mathbf{W}^{-1} \mathbf{U}. \quad (42)$$

The projection of these K eigenvectors to model space span the special subspace and so the L columns of \mathbf{U} can be used to define the Householder matrix \mathbf{X} . Recall the two bullet points at the end of Sec. 5, where $\delta \mathbf{x}_s$ is a vector that spans this special space. Appendix C shows how \mathbf{X} can be constructed from \mathbf{U} .

The aim of the following procedure is to use information from the EMs to define the first K columns of $\mathbf{P}_\chi^{f^{-1}}$ as defined in (18) and used in (19). In a similar fashion to the HSV approach in Sec. 6, let

$$\mathbf{Z} = \mathbf{P}^{f^{-1}} \mathbf{U}, \quad (43)$$

$$= \mathbf{P}^{f^{-1}} \mathbf{L} \mathbf{X} \mathbf{X}^T \mathbf{L}^{-1} \mathbf{U},$$

$$\mathbf{X}^T \mathbf{L}^T \mathbf{Z} = \mathbf{X}^T \mathbf{L}^T \mathbf{P}^{f^{-1}} \mathbf{L} \mathbf{X} \mathbf{X}^T \mathbf{L}^{-1} \mathbf{U},$$

$$= \mathbf{P}_\chi^{f^{-1}} \hat{\mathbf{I}} \mathbf{X}^T \mathbf{L}^{-1} \mathbf{U}, \quad (44)$$

where $\hat{\mathbf{I}}$ is the $K \times N$ non-square quasi-identity matrix defined by (24). This has to be included because $\mathbf{P}_\chi^{f^{-1}}$ is defined only with the first K columns. Equation (44) leads to

$$\mathbf{P}_\chi^{f^{-1}} = \mathbf{X}^T \mathbf{L}^T \mathbf{Z} (\hat{\mathbf{I}} \mathbf{X}^T \mathbf{L}^{-1} \mathbf{U})^{-1}, \quad (45)$$

where the operator inverted is a calculable $K \times K$ matrix, which we assume is non-singular. Note that (45) is for only part of the inverse covariance matrix and so is not symmetric. The matrix yet known is $\mathbf{X}^T \mathbf{L}^T \mathbf{Z}$ which is now found from the EMs (Sec. 3).

The \mathbf{P}^f that is of interest exists only in the K -dimensional subspace. The \mathbf{P}^f that exists in the L -dimensional subspace is (35). This can be modified to exist in the K -dimensional subspace replacing \mathbf{P}_L^f in (35) with \mathbf{P}_K^f found from (40).

$$\mathbf{P}^f = \tilde{\mathbf{S}} \mathbf{P}_L^f \tilde{\mathbf{S}}^T \rightarrow \tilde{\mathbf{S}} \mathbf{P}_K^f \tilde{\mathbf{S}}^T = \tilde{\mathbf{S}} \mathbf{U}_K \Lambda_K \mathbf{U}_K^T \tilde{\mathbf{S}}^T. \quad (46)$$

Equation (46) is developed as follows (steps marked with a * need special note - see below)

$$\mathbf{I} = \tilde{\mathbf{S}} \mathbf{U}_K \Lambda_K \mathbf{U}_K^T \tilde{\mathbf{S}}^T \mathbf{P}^{f^{-1}},$$

$$\tilde{\mathbf{S}}^T \mathbf{W}^{-1} = \mathbf{U}_K \Lambda_K \mathbf{U}_K^T \tilde{\mathbf{S}}^T \mathbf{P}^{f^{-1}},$$

$$\mathbf{U}_K^T \tilde{\mathbf{S}}^T \mathbf{W}^{-1} = \Lambda_K \mathbf{U}_K^T \tilde{\mathbf{S}}^T \mathbf{P}^{f^{-1}},$$

$$\Lambda_K^{-1} \mathbf{U}_K^T \tilde{\mathbf{S}}^T \mathbf{W}^{-1} = \mathbf{U}_K^T \tilde{\mathbf{S}}^T \mathbf{P}^{f^{-1}},$$

$$\mathbf{U}_K \Lambda_K^{-1} \mathbf{U}_K^T \tilde{\mathbf{S}}^T \mathbf{W}^{-1} = \tilde{\mathbf{S}}^T \mathbf{P}^{f^{-1}}. \quad (47 *)$$

The following, derived from (7) and (39) holds as long as it acts only on vectors entirely in the column space of $\tilde{\mathbf{S}}$

$$\tilde{\mathbf{S}} \tilde{\mathbf{S}}^T \mathbf{W}^{-1} = \mathbf{I}. \quad (48 *)$$

When used with (47) this gives

$$\mathbf{U}_K \Lambda_K^{-1} \mathbf{U}_K^T \tilde{\mathbf{S}}^T \mathbf{W}^{-1} = \tilde{\mathbf{S}}^T \mathbf{W}^{-1} \mathbf{W} \mathbf{P}^{f^{-1}},$$

$$\tilde{\mathbf{S}} \mathbf{U}_K \Lambda_K^{-1} \mathbf{U}_K^T \tilde{\mathbf{S}}^T \mathbf{W}^{-1} = \mathbf{W} \mathbf{P}^{f^{-1}},$$

$$\mathbf{W}^{-1} \tilde{\mathbf{S}} \mathbf{U}_K \Lambda_K^{-1} \mathbf{U}_K^T \tilde{\mathbf{S}}^T \mathbf{W}^{-1} = \mathbf{P}^{f^{-1}}. \quad (49)$$

The steps marked with a (*) require special attention as these statements are appropriate only under special circumstances. This may require some checking. Equation (49) is used in the definition of \mathbf{Z} (43), which is itself used for the definition of $\mathbf{P}_\chi^{\text{f-1}}$ in (45). This gives

$$\mathbf{P}_\chi^{\text{f-1}} = \mathbf{X}^T \mathbf{L}^T \mathbf{W}^{-1} \tilde{\mathbf{S}} \mathbf{U}_K \Lambda_K^{-1} \mathbf{U}_K^T \tilde{\mathbf{S}}^T \mathbf{W}^{-1} \mathbf{U} (\hat{\mathbf{I}} \mathbf{X}^T \mathbf{L}^{-1} \mathbf{U})^{-1}, \quad (50)$$

which is a $N \times K$ matrix as required.

8. Differentiating the cost function

The background part of the cost function, J_b (19), is now defined, where $\mathbf{P}_\chi^{\text{f-1}}$ is found depending upon whether the HSV or the EM source of information is used. For the HSV calculation, $\mathbf{P}_\chi^{\text{f-1}}$ is found from (30) and for the EM calculation, $\mathbf{P}_\chi^{\text{f-1}}$ is found from (50). For VAR, the derivatives of J_b and J_o (the observation part of the cost function) are required with respect to each component of the control vector χ . Recall that the control vector comprises two parts: one that describes the special subspace, χ_s (non-zero only in the first K components), and another that describes the remainder, $\bar{\chi}_s$ (non-zero only in the last $N - K$ components)

$$\chi = \chi_s + \bar{\chi}_s, \quad (51)$$

and the gradient vector, $\partial J / \partial \chi$ has a similar structure (i.e. $\partial J / \partial \chi_s$ is non-zero only in the first K components and $\partial J / \partial \bar{\chi}_s$ is non-zero only in the last $N - K$ components). J is the total cost function, which is the sum of the background part, J_b in (19) and an observation part, J_o . Let the three terms defining J_b in (19) be written separately

$$J = J_b^{\text{I}} + J_b^{\text{II}} + J_b^{\text{III}} + J_o, \quad (52)$$

$$\text{where } J_b^{\text{I}} = \frac{1}{2} (\chi_s - \chi_s^{\text{b}})^T \mathbf{P}_\chi^{\text{f-1}} (\chi_s - \chi_s^{\text{b}}), \quad (53)$$

$$J_b^{\text{II}} = \alpha (\bar{\chi}_s - \bar{\chi}_s^{\text{b}})^T \mathbf{P}_\chi^{\text{f-1}} (\chi_s - \chi_s^{\text{b}}), \quad (54)$$

$$\text{and } J_b^{\text{III}} = \frac{1}{2} (\bar{\chi}_s - \bar{\chi}_s^{\text{b}})^T (\bar{\chi}_s - \bar{\chi}_s^{\text{b}}). \quad (55)$$

It is assumed that $\partial J_o / \partial \delta \mathbf{x}$ has already been found, e.g. by the conventional adjoint method (e.g. Rodgers, 200x; Bannister 200x). There are eight contributions to the gradient vector, viz. $\partial J_b^{\text{I}} / \partial \chi_s$, $\partial J_b^{\text{I}} / \partial \bar{\chi}_s$, $\partial J_b^{\text{II}} / \partial \chi_s$, $\partial J_b^{\text{II}} / \partial \bar{\chi}_s$, $\partial J_b^{\text{III}} / \partial \chi_s$, $\partial J_b^{\text{III}} / \partial \bar{\chi}_s$, $\partial J_o / \partial \chi_s$, and $\partial J_o / \partial \bar{\chi}_s$, where each contributes to the total gradient vector in the following way

$$\frac{\partial J}{\partial \chi} = \frac{\partial J_b^{\text{I}}}{\partial \chi_s} + \frac{\partial J_b^{\text{I}}}{\partial \bar{\chi}_s} + \frac{\partial J_b^{\text{II}}}{\partial \chi_s} + \frac{\partial J_b^{\text{II}}}{\partial \bar{\chi}_s} + \frac{\partial J_b^{\text{III}}}{\partial \chi_s} + \frac{\partial J_b^{\text{III}}}{\partial \bar{\chi}_s} + \frac{\partial J_o}{\partial \chi_s} + \frac{\partial J_o}{\partial \bar{\chi}_s}. \quad (56)$$

Note that derivatives with respect to χ_s give a vector that is non-zero only in the first K components, and so imply differentiating with respect to the first K components of χ only. Similarly derivatives with respect to $\bar{\chi}_s$ give a vector that is non-zero only in the last $N - K$ components, and so imply differentiating with respect to the last $N - K$ components of χ only.

It will be useful to expand-out the matrix notation in (53)-(55) as follows (in the following, ignore the χ_s^{b} and $\bar{\chi}_s^{\text{b}}$ terms for simplicity (they can be put back in later with no loss of generality). Whether we are interested in χ_s or $\bar{\chi}_s$ will be controlled by the range of the index (1 to K or $K + 1$ to N respectively)

$$J_b^{\text{I}} = \frac{1}{2} \sum_{i=1}^K (\chi_s)_i \sum_{j=1}^K (\mathbf{P}_\chi^{\text{f-1}})_{ij} (\chi_s)_j = \frac{1}{2} \sum_{i=1}^K \chi_i \sum_{j=1}^K (\mathbf{P}_\chi^{\text{f-1}})_{ij} \chi_j, \quad (57)$$

$$J_b^{\text{II}} = \alpha \sum_{i=K+1}^N (\bar{\chi}_s)_i \sum_{j=1}^K (\mathbf{P}_\chi^{\text{r-1}})_{ij} (\chi_s)_j = \alpha \sum_{i=K+1}^N \chi_i \sum_{j=1}^K (\mathbf{P}_\chi^{\text{r-1}})_{ij} \chi_j, \quad (58)$$

$$J_b^{\text{III}} = \frac{1}{2} \sum_{i=K+1}^N \chi_i^2. \quad (59)$$

Each contribution in (56) is now addressed in turn.

Contribution (i): $\partial J_b^{\text{I}} / \partial \chi_s$

Differentiating (57) with respect to χ (components $1 \leq k \leq K$) gives

$$\begin{aligned} \frac{\partial J_b^{\text{I}}}{\partial (\chi_s)_k} &= \frac{1}{2} \sum_{i=1}^K \delta_{ik} \sum_{j=1}^K (\mathbf{P}_\chi^{\text{r-1}})_{ij} \chi_j + \frac{1}{2} \sum_{i=1}^K \chi_i \sum_{j=1}^K (\mathbf{P}_\chi^{\text{r-1}})_{ij} \delta_{jk}, \\ &= \frac{1}{2} \sum_{j=1}^K (\mathbf{P}_\chi^{\text{r-1}})_{kj} \chi_j + \frac{1}{2} \sum_{i=1}^K \chi_i (\mathbf{P}_\chi^{\text{r-1}})_{ik} = \sum_{j=1}^K (\mathbf{P}_\chi^{\text{r-1}})_{kj} \chi_j. \end{aligned} \quad (60)$$

Contribution (ii): $\partial J_b^{\text{I}} / \partial \bar{\chi}_s$

Differentiating (57) with respect to χ (components $K+1 \leq k \leq N$) gives

$$\frac{\partial J_b^{\text{I}}}{\partial (\bar{\chi}_s)_k} = 0. \quad (61)$$

Contribution (iii): $\partial J_b^{\text{II}} / \partial \chi_s$

Differentiating (58) with respect to χ (components $1 \leq k \leq K$) gives

$$\begin{aligned} \frac{\partial J_b^{\text{II}}}{\partial (\chi_s)_k} &= \alpha \sum_{i=K+1}^N \chi_i \sum_{j=1}^K (\mathbf{P}_\chi^{\text{r-1}})_{ij} \delta_{jk}, \\ &= \alpha \sum_{i=K+1}^N \chi_i (\mathbf{P}_\chi^{\text{r-1}})_{ik}. \end{aligned} \quad (62)$$

Contribution (iv): $\partial J_b^{\text{II}} / \partial \bar{\chi}_s$

Differentiating (58) with respect to χ (components $K+1 \leq k \leq N$) gives

$$\begin{aligned} \frac{\partial J_b^{\text{II}}}{\partial (\bar{\chi}_s)_k} &= \alpha \sum_{i=K+1}^N \delta_{ik} \sum_{j=1}^K (\mathbf{P}_\chi^{\text{r-1}})_{ij} \chi_j, \\ &= \alpha \sum_{j=1}^K (\mathbf{P}_\chi^{\text{r-1}})_{kj} \chi_j. \end{aligned} \quad (63)$$

Contribution (v): $\partial J_b^{\text{III}} / \partial \chi_s$

Differentiating (59) with respect to χ (components $1 \leq k \leq K$) gives

$$\frac{\partial J_b^{\text{III}}}{\partial (\chi_s)_k} = 0. \quad (64)$$

Contribution (vi): $\partial J_b^{\text{III}} / \partial \bar{\chi}_s$

Differentiating (59) with respect to χ (components $N+1 \leq k \leq N$) gives

$$\begin{aligned} \frac{\partial J_b^{\text{III}}}{\partial (\chi_s)_k} &= \sum_{i=K+1}^N \chi_i \delta_{ik}, \\ &= \chi_k. \end{aligned} \quad (65)$$

Contributions (vii) and (viii): $\partial J_o / \partial \chi_s$ and $\partial J_o / \partial \bar{\chi}_s$

For this term it is not necessary to distinguish between parts of the state vector. Using the definition of the control variable transform in (13), it can be expanded as

$$\delta \mathbf{x}_m = \sum_{n=1}^N \mathbf{L}_{mn} \sum_{p=1}^N \mathbf{X}_{np} \chi_p. \quad (66)$$

Therefore

$$\frac{\partial \mathbf{x}_m}{\partial \chi_q} = \sum_{n=1}^N \mathbf{L}_{mn} \sum_{p=1}^N \mathbf{X}_{np} \delta_{pq} = \sum_{n=1}^N \mathbf{L}_{mn} \mathbf{X}_{nq}. \quad (67)$$

Using the chain rule, and then feeding-in (67) gives

$$\begin{aligned} \frac{\partial J_o}{\partial \chi_q} &= \sum_{m=1}^N \frac{\partial \mathbf{x}_m}{\partial \chi_q} \frac{\partial J_o}{\partial \delta \mathbf{x}_m}, \\ &= \sum_{m=1}^N \sum_{n=1}^N \mathbf{L}_{mn} \mathbf{X}_{nq} \frac{\partial J_o}{\partial \delta \mathbf{x}_m}, \\ &= \sum_{n=1}^N \mathbf{X}_{qn}^T \sum_{m=1}^N \mathbf{L}_{nm}^T \frac{\partial J_o}{\partial \delta \mathbf{x}_m}. \end{aligned} \quad (68)$$

Equation (68) is equivalent to the matrix operation represented by

$$\frac{\partial J_o}{\partial \chi} = \mathbf{X}^T \mathbf{L}^T \frac{\partial J_o}{\partial \delta \mathbf{x}}, \quad (69)$$

which is the standard result of the adjoint of the control variable transform.

References

Bannister R.N., 2001, Elementary 4d-Var, DARC Technical Report No. 2, 16 Pp.

Bannister R.N., 2008, A review of forecast error covariance statistics in atmospheric variational data assimilation. I: Characteristics and measurements of forecast error covariances, Q.J.R. Meteorol. Soc. 134, 1951-1970.

Fisher M., 1998, Development of a simplified Kalman filter, ECMWF Tech. Memo. 260 [available from ECMWF, Shinfield Park, Reading, Berkshire, RG2 9AX, UK].

Fisher M., 2007, The sensitivity of analysis errors to the specification of background error covariances, Pp. 27-36 of ECMWF Workshop Proceedings on Flow-dependent aspects of Data Assimilation, 11-13 June 2007, ECMWF, Reading, UK.

Press W.H., Teukolsky S.A., Vetterling W.T., Flannery B.P. 1986, Numerical Recipes in Fortran, Cambridge University Press.

Rodgers C.D., Inverse methods for atmospheric sounding: theory and practice (Sec. 8.2.3). Series on atmospheric, oceanic and planetary physics, Vol. 2, World Scientific.

Appendix A: the Gram-Schmidt procedure

The creation of an orthogonalized set of vectors, \tilde{s}_i , from a set that is made up of non-orthogonal vectors, s_i , may be performed by the Gram-Schmidt procedure. The space spanned by each set of vectors is the same, but the orthogonalized set is more convenient to work with. Let orthogonalization be performed with respect to the \mathbf{W}^{-1} inner product.

Let the first vector of the orthogonalized set, \tilde{s}_1 , be the first vector of the non-orthogonal set, s_1 , but normalized to have unit length (under the \mathbf{W}^{-1} norm).

$$\tilde{s}_1 = \frac{1}{N_1} s_1, \quad (\text{A.1})$$

$$\tilde{s}_1^T \mathbf{W}^{-1} \tilde{s}_1 = 1, \quad (\text{A.2})$$

$$\therefore N_1 = \sqrt{s_1^T \mathbf{W}^{-1} s_1}. \quad (\text{A.3})$$

For $i \geq 1$, the $i+1$ th orthogonalized vector is defined as the $i+1$ th non-orthogonal vector minus a linear combination of all previous vectors defined (and normalized)

$$\tilde{s}_{i+1} = \frac{1}{N_{i+1}} \left(s_{i+1} - \sum_{j=1}^i \alpha_{j,i+1} \tilde{s}_j \right). \quad (\text{A.4})$$

The coefficients, $\alpha_{j,i+1}$, are chosen for orthogonality as follows

$$\tilde{s}_j^T \mathbf{W}^{-1} \tilde{s}_{i+1} = \delta_{j,i+1}. \quad (\text{A.5})$$

Performing an inner product of (A.4) with \tilde{s}_k (where $1 \leq k \leq i$) gives

$$\tilde{s}_k^T \mathbf{W}^{-1} \tilde{s}_{i+1} = \frac{1}{N_{i+1}} \left(\tilde{s}_k^T \mathbf{W}^{-1} s_{i+1} - \sum_{j=1}^i \alpha_{j,i+1} \tilde{s}_k^T \mathbf{W}^{-1} \tilde{s}_j \right) = 0. \quad (\text{A.6})$$

This is set equal to zero by (A.5) and by the fact that $k \neq i+1$. Further use of (A.5) leads to

$$\begin{aligned} \tilde{s}_k^T \mathbf{W}^{-1} s_{i+1} - \alpha_{k,i+1} &= 0, \\ \therefore \alpha_{k,i+1} &= \tilde{s}_k^T \mathbf{W}^{-1} s_{i+1}. \end{aligned} \quad (\text{A.7})$$

This determines the coefficients. Let \tilde{t}_{i+1} be the part of (A.4) inside the brackets (i.e. the unnormalized vector)

$$\tilde{s}_{i+1} = \frac{1}{N_{i+1}} \tilde{t}_{i+1}, \quad (\text{A.8})$$

$$\text{where } \tilde{t}_{i+1} = s_{i+1} - \sum_{j=1}^i \alpha_{j,i+1} \tilde{s}_j, \quad (\text{A.9})$$

which can be calculated now that the coefficients are known (A.7). N_{i+1} then follows in a similar way to (A.2) and (A.3)

$$\tilde{s}_{i+1}^T \mathbf{W}^{-1} \tilde{s}_{i+1} = 1, \quad (\text{A.10})$$

$$\therefore N_{i+1} = \sqrt{\tilde{t}_{i+1}^T \mathbf{W}^{-1} \tilde{t}_{i+1}}. \quad (\text{A.11})$$

Appendix B: Proof of the orthogonality of the residual in (9)

In these notes a vector in model space, $\delta \mathbf{x}$ is divided into a part that is spanned by columns of $\tilde{\mathbf{S}}$, $\delta \mathbf{x}_s$, and a residual, $\delta \tilde{\mathbf{x}}_s$, given by (9). Here it is proved that any vector that exists only in the

spanned space is ' \mathbf{W}^{-1} -orthogonal' to one that exists only in the residual. Mathematically, it is to be shown that

$$\delta\bar{\mathbf{x}}_s^T \mathbf{W}^{-1} \delta\mathbf{x}_s = 0. \quad (\text{B.1})$$

First eliminate $\delta\bar{\mathbf{x}}_s$ using (9)

$$\delta\bar{\mathbf{x}}_s^T \mathbf{W}^{-1} \delta\mathbf{x}_s = (\delta\mathbf{x} - \delta\mathbf{x}_s)^T \mathbf{W}^{-1} \delta\mathbf{x}_s. \quad (\text{B.2})$$

$\delta\mathbf{x}_s$ can be written in terms of $\delta\mathbf{x}$ by combining (6) and (8)

$$\delta\mathbf{x}_s = \tilde{\mathbf{S}}\tilde{\mathbf{S}}^T \mathbf{W}^{-1} \delta\mathbf{x}. \quad (\text{B.3})$$

Substituting (B.3) into (B.2) gives

$$\begin{aligned} \delta\bar{\mathbf{x}}_s^T \mathbf{W}^{-1} \delta\mathbf{x}_s &= (\delta\mathbf{x} - \tilde{\mathbf{S}}\tilde{\mathbf{S}}^T \mathbf{W}^{-1} \delta\mathbf{x})^T \mathbf{W}^{-1} \tilde{\mathbf{S}}\tilde{\mathbf{S}}^T \mathbf{W}^{-1} \delta\mathbf{x} \\ &= \delta\mathbf{x}^T \mathbf{W}^{-1} \tilde{\mathbf{S}}\tilde{\mathbf{S}}^T \mathbf{W}^{-1} \delta\mathbf{x} - \delta\mathbf{x}^T \mathbf{W}^{-1} \tilde{\mathbf{S}}\tilde{\mathbf{S}}^T \mathbf{W}^{-1} \tilde{\mathbf{S}}\tilde{\mathbf{S}}^T \mathbf{W}^{-1} \delta\mathbf{x}. \end{aligned} \quad (\text{B.4})$$

Next use the orthogonality of the subspace (7) which proves that each term in (B.4) is equal and opposite, thus giving zero and proving (B.1)

$$\delta\bar{\mathbf{x}}_s^T \mathbf{W}^{-1} \delta\mathbf{x}_s = \delta\mathbf{x}^T \mathbf{W}^{-1} \tilde{\mathbf{S}}\tilde{\mathbf{S}}^T \mathbf{W}^{-1} \delta\mathbf{x} - \delta\mathbf{x}^T \mathbf{W}^{-1} \tilde{\mathbf{S}}\tilde{\mathbf{S}}^T \mathbf{W}^{-1} \delta\mathbf{x} = 0. \quad (\text{B.5})$$

Appendix C: Design of the sequence of Householder transforms

It is now shown how \mathbf{X} can be formulated to achieve property (20). Fisher (1998) states that this can be achieved with a sequence of Householder transformations. A single Householder transformation, \mathbf{P} , (e.g. Press et al. 1986) may be written as follows

$$\mathbf{P} = \mathbf{I} - 2 \frac{\mathbf{u}\mathbf{u}^T}{\mathbf{u}^T \mathbf{u}}, \quad (\text{C.1})$$

where

$$\mathbf{u} = \mathbf{x} \mp |\mathbf{x}| \mathbf{e}_1. \quad (\text{C.2})$$

The vector \mathbf{x} is arbitrary and \mathbf{e}_1 is a vector which is zero valued except for the first element, which has unit value (the properties of the Householder transformation hold for a general single element being chosen instead, although here we always choose the first element). \mathbf{P} is useful because it has the following useful properties.

- The Householder transformation is orthogonal

$$\begin{aligned} \mathbf{P}\mathbf{P}^T &= \left(\mathbf{I} - 2 \frac{\mathbf{u}\mathbf{u}^T}{\mathbf{u}^T \mathbf{u}} \right) \left(\mathbf{I} - 2 \frac{\mathbf{u}\mathbf{u}^T}{\mathbf{u}^T \mathbf{u}} \right)^T, \\ &= \mathbf{I} - 4 \frac{\mathbf{u}\mathbf{u}^T}{\mathbf{u}^T \mathbf{u}} + 4 \frac{\mathbf{u}\mathbf{u}^T \mathbf{u}\mathbf{u}^T}{\mathbf{u}^T \mathbf{u}\mathbf{u}^T \mathbf{u}}, \\ &= \mathbf{I} - 4 \frac{\mathbf{u}\mathbf{u}^T}{\mathbf{u}^T \mathbf{u}} + 4 \frac{\mathbf{u}\mathbf{u}^T}{\mathbf{u}^T \mathbf{u}} = \mathbf{I}. \end{aligned} \quad (\text{C.3})$$

- When acting on the state \mathbf{x} , which is used to define \mathbf{P} in (C.1) and (C.2), the result is a vector with all but the first element zero

$$\begin{aligned} \mathbf{P}\mathbf{x} &= \left(\mathbf{I} - 2 \frac{\mathbf{u}\mathbf{u}^T}{\mathbf{u}^T \mathbf{u}} \right) \mathbf{x}, \\ &= \left(\mathbf{I} - 2 \frac{\mathbf{u}(\mathbf{x} \mp |\mathbf{x}| \mathbf{e}_1)^T}{\mathbf{u}^T \mathbf{u}} \right) \mathbf{x}, \end{aligned}$$

$$\begin{aligned}
 &= \left(\mathbf{I} - 2 \frac{\mathbf{u}(\mathbf{x} \mp |\mathbf{x}| \mathbf{e}_1)^T}{2\mathbf{x}^T\mathbf{x} \mp 2|\mathbf{x}|x_1} \right) \mathbf{x}, \\
 &= \mathbf{x} - 2 \frac{\mathbf{u}(\mathbf{x}^T\mathbf{x} \mp |\mathbf{x}|x_1)}{2\mathbf{x}^T\mathbf{x} \mp 2|\mathbf{x}|x_1}, \\
 &= \mathbf{x} - \mathbf{u}, \\
 &= \pm |\mathbf{x}| \mathbf{e}_1.
 \end{aligned} \tag{C.4}$$

- When acting on a state $\bar{\mathbf{x}}$, which is orthogonal to the state \mathbf{x} , which is used to define \mathbf{P} in (C.1) and (C.2), the result is a vector with zero in the first element

$$\begin{aligned}
 \text{Note } \mathbf{u}^T \bar{\mathbf{x}} &= \mathbf{x}^T \bar{\mathbf{x}} \mp |\mathbf{x}| \mathbf{e}_1^T \bar{\mathbf{x}}, \\
 &= \mp |\mathbf{x}| \bar{x}_1,
 \end{aligned} \tag{C.5}$$

$$\begin{aligned}
 \text{then } \mathbf{P}\bar{\mathbf{x}} &= \left(\mathbf{I} - 2 \frac{\mathbf{u}\mathbf{u}^T}{\mathbf{u}^T\mathbf{u}} \right) \bar{\mathbf{x}}, \\
 &= \bar{\mathbf{x}} \pm 2 \frac{|\mathbf{x}| \bar{x}_1 (\mathbf{x} \mp |\mathbf{x}| \mathbf{e}_1)}{2\mathbf{x}^T\mathbf{x} \mp 2|\mathbf{x}|x_1}, \\
 &= \frac{|\mathbf{x}|^2 \bar{\mathbf{x}} \mp |\mathbf{x}|x_1 \bar{\mathbf{x}} \pm |\mathbf{x}| \bar{x}_1 \mathbf{x} - \bar{x}_1 |\mathbf{x}|^2 \mathbf{e}_1}{\mathbf{x}^T\mathbf{x} \mp |\mathbf{x}|x_1}.
 \end{aligned} \tag{C.6}$$

Equation (C.6) does not have weight in element 1. To show this, do a scalar product with \mathbf{e}_1

$$\mathbf{e}_1^T \mathbf{P}\bar{\mathbf{x}} = \frac{|\mathbf{x}|^2 \bar{x}_1 \mp |\mathbf{x}|x_1 \bar{x}_1 \pm |\mathbf{x}| \bar{x}_1 x_1 - \bar{x}_1 |\mathbf{x}|^2}{\mathbf{x}^T\mathbf{x} \mp |\mathbf{x}|x_1} = 0, \tag{C.7}$$

In these equations, \bar{x}_1 and x_1 are the first components of $\bar{\mathbf{x}}$ and \mathbf{x} respectively. The first property gives $\mathbf{P} = \mathbf{P}^T = \mathbf{P}^{-1}$. These properties can be combined to give \mathbf{X} in the following way. Defining $\mathbf{R}^{(0)} = \mathbf{L}^{-1}\mathbf{S}$ for the HSV calculation or $\mathbf{R}^{(0)} = \mathbf{L}^{-1}\mathbf{U}$ for the EM calculation, let $\mathbf{X}^T \mathbf{R}^{(0)}$ be a vector of two parts

$$\mathbf{X}^T \mathbf{R}^{(0)} = \begin{pmatrix} \mathbf{A} \\ \mathbf{0} \end{pmatrix}, \tag{C.8}$$

where \mathbf{A} is a $K \times K$ matrix consistent with the required property of (20). In fact, by the way that \mathbf{X}^T is to be formed, \mathbf{A} will turn out to be upper triangular. Let

$$\mathbf{X}^T \mathbf{R}^{(0)} = \mathbf{P}_K \dots \mathbf{P}_k \dots \mathbf{P}_2 \mathbf{P}_1 \mathbf{R}^{(0)}. \tag{C.9}$$

Each \mathbf{P}_i transformation is Householder-like according to the following, e.g. for \mathbf{P}_1

$$\mathbf{P}_1 \mathbf{R}^{(0)} = \left(\mathbf{I} - 2 \frac{\mathbf{u}\mathbf{u}^T}{\mathbf{u}^T\mathbf{u}} \right) \mathbf{R}^{(0)}, \quad \mathbf{u} = \mathbf{r}_1^{(0)} - |\mathbf{r}_1^{(0)}| \mathbf{e}_1^{(N)}, \tag{C.10}$$

where $\mathbf{r}_1^{(0)}$ is the first column of $\mathbf{R}^{(0)}$ and $\mathbf{e}_1^{(N)}$ is the N -element vector with all but the first element zero (which is unity). This generates a new matrix $\mathbf{R}^{(1)} = \mathbf{P}_1 \mathbf{R}^{(0)}$ which has the form

$$\mathbf{R}^{(1)} = \begin{pmatrix} r_{11}^{(1)} & r_{12}^{(1)} & r_{13}^{(1)} & \dots & r_{1K}^{(1)} \\ 0 & r_{22}^{(1)} & r_{23}^{(1)} & \dots & r_{2K}^{(1)} \\ 0 & r_{32}^{(1)} & r_{33}^{(1)} & \dots & r_{3K}^{(1)} \\ \dots & \dots & \dots & \dots & \dots \\ 0 & r_{k2}^{(1)} & r_{k3}^{(1)} & \dots & r_{kK}^{(1)} \\ \dots & \dots & \dots & \dots & \dots \\ 0 & r_{N2}^{(1)} & r_{N3}^{(1)} & \dots & r_{NK}^{(1)} \end{pmatrix}, \tag{C.11}$$

having only the first element non-zero of the first column (since \mathbf{P}_1 is designed in terms of the first column of $\mathbf{R}^{(0)}$). The aim now is to act with a $N - 1 \times N - 1$ element Householder operator on $\mathbf{R}^{(1)}$ excluding the first row and first column

$$\mathbf{P}_2 \mathbf{R}^{(1)} = \begin{pmatrix} 1 & \mathbf{0} \\ \mathbf{0} & \mathbf{I} - 2 \frac{\mathbf{u}\mathbf{u}^T}{\mathbf{u}^T \mathbf{u}} \end{pmatrix} \mathbf{R}^{(1)}, \quad \mathbf{u} = \mathbf{r}_2^{(1)} - |\mathbf{r}_2^{(1)}| \mathbf{e}_1^{(N-1)}, \quad (\text{C.12})$$

where the partitioned-off (lower right) part of \mathbf{P}_2 is a $N - 1 \times N - 1$ matrix, $\mathbf{r}_2^{(1)}$ is the $N - 1$ -element second column of $\mathbf{R}^{(1)}$ (excluding the first component), and $\mathbf{e}_1^{(N-1)}$ is the $N - 1$ -element vector with all but the first element zero (which is unity). This generates a new matrix $\mathbf{R}^{(2)} = \mathbf{P}_2 \mathbf{R}^{(1)}$ which has the form

$$\mathbf{R}^{(2)} = \begin{pmatrix} r_{11}^{(1)} & r_{12}^{(1)} & r_{13}^{(1)} & \dots & r_{1K}^{(1)} \\ 0 & r_{22}^{(2)} & r_{23}^{(2)} & \dots & r_{2K}^{(2)} \\ 0 & 0 & r_{33}^{(2)} & \dots & r_{3K}^{(2)} \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & r_{k3}^{(2)} & \dots & r_{kK}^{(2)} \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & r_{N3}^{(2)} & \dots & r_{NK}^{(2)} \end{pmatrix}. \quad (\text{C.13})$$

The k th operator, \mathbf{P}_k , has the form

$$\mathbf{P}_k \mathbf{R}^{(k-1)} = \begin{pmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{0} & \mathbf{I} - 2 \frac{\mathbf{u}\mathbf{u}^T}{\mathbf{u}^T \mathbf{u}} \end{pmatrix} \mathbf{R}^{(k-1)}, \quad \mathbf{u} = \mathbf{r}_k^{(k-1)} - |\mathbf{r}_k^{(k-1)}| \mathbf{e}_1^{(N-k+1)}, \quad (\text{C.14})$$

where the partitioned-off part of \mathbf{P}_k is a $N - k + 1 \times N - k + 1$ matrix, $\mathbf{r}_k^{(k-1)}$ is the $N - k + 1$ -element k th column of $\mathbf{R}^{(k-1)}$ (excluding the first $k - 1$ components) and $\mathbf{e}_1^{(N-k+1)}$ is the $N - k + 1$ -element vector with all but the first element zero (which is unity). After all K operators have acted, the result is $\mathbf{R}^{(k)} = \mathbf{X}^T \mathbf{R}^{(0)}$

$$\mathbf{R}^{(k)} = \begin{pmatrix} r_{11}^{(1)} & r_{12}^{(1)} & r_{13}^{(1)} & \dots & r_{1K}^{(1)} \\ 0 & r_{22}^{(2)} & r_{23}^{(2)} & \dots & r_{2K}^{(2)} \\ 0 & 0 & r_{33}^{(3)} & \dots & r_{3K}^{(3)} \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & r_{KK}^{(k)} \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & 0 \end{pmatrix}. \quad (\text{C.15})$$

where the top section is the matrix \mathbf{A} in (C.8), and the bottom section comprises zeros.

It should also be shown that $\mathbf{X}\mathbf{X}^T = \mathbf{I}$. From (C.9) this is easy to show given that each pair has the property that $\mathbf{P}_k \mathbf{P}_k^T = \mathbf{I}$

$$\mathbf{P}_K \dots \mathbf{P}_k \dots \mathbf{P}_2 \mathbf{P}_1 \mathbf{P}_1^T \mathbf{P}_2^T \dots \mathbf{P}_k^T \dots \mathbf{P}_K^T = \mathbf{I}. \quad (\text{C.16})$$

It remains to be shown that the string of Householder operators $\mathbf{P}_K \dots \mathbf{P}_k \dots \mathbf{P}_2 \mathbf{P}_1$ acting on a state, $\bar{\mathbf{r}}^{(0)}$ (which is orthogonal to all columns of $\mathbf{R}^{(0)}$) gives a state that is zero in the first K elements. All \mathbf{P}_k are formed in the same way as shown above (ie with respect to the $\mathbf{R}^{(k)}$ matrices).

First let $\bar{\mathbf{r}}^{(1)} = \mathbf{P}_1 \bar{\mathbf{r}}^{(0)}$. Since $\bar{\mathbf{r}}^{(0)}$ is orthogonal to $\mathbf{r}_1^{(0)}$ (the latter is the vector used to define \mathbf{P}_1 in (C.10)), and by property (C.7), the vector $\bar{\mathbf{r}}^{(1)}$ has zero in the first element. Next, let

$\bar{\mathbf{r}}^{(2)} = \mathbf{P}_2 \bar{\mathbf{r}}^{(1)}$. By a similar argument, if the vector formed from the last $N - 1$ components of $\bar{\mathbf{r}}^{(1)}$ is orthogonal to the vector formed from the last $N - 1$ components of $\mathbf{r}_2^{(1)}$ (the latter is the vector used to define \mathbf{P}_2 in (C.12)), and by property (C.7), the vector $\bar{\mathbf{r}}^{(2)}$ will have zero in the first two elements. Because the first element of $\mathbf{P}_1 \bar{\mathbf{r}}^{(0)}$ is always zero, the remaining $N - 1$ -component inner product in question is equal to the full N -component inner product as the first element contributes zero. The orthogonality test is therefore satisfied if the following N -component inner product is zero

$$(\mathbf{P}_1 \bar{\mathbf{r}}^{(0)})^T \mathbf{P}_1 \mathbf{r}_2^{(0)} = \bar{\mathbf{r}}^{(0)T} \mathbf{P}_1^T \mathbf{P}_1 \mathbf{r}_2^{(0)} = \bar{\mathbf{r}}^{(0)T} \mathbf{r}_2^{(0)} = 0. \quad (\text{C.17})$$

This is satisfied because the vector $\bar{\mathbf{r}}^{(0)}$ is orthogonal to $\mathbf{r}_2^{(0)}$ by definition. These arguments continue for all K operators. The final result is a vector of zeros in the first K components.