ABC Vn1.4da

April 26, 2018

1 Subroutine convention

1.1 Input and output in systematic subroutines

- Forward parameter order: (LS), input structure, output structure, ...
- Inverse parameter order: (LS), output structure, input structure, ...
- Adjoint parameter order: (LS), output structure, input structure, ...

2 Parameter transform

The control variables are (1) streamfunction ($\delta\psi$), (2) velocity potential ($\delta\chi$), (3) unbalanced $\delta\tilde{\rho}'$ ($\delta^{\mathrm{u}}\tilde{\rho}'$), (4) unbalanced b' ($\delta^{\mathrm{u}}b'$), (5) unbalanced w ($\delta^{\mathrm{u}}w$), and (6) tracer (δq). Here is the procedure for the standard transform $\mathbf{U}_{\mathrm{p}}$.

1. Compute u and v:
$$\delta u = \frac{\partial \delta\chi}{\partial x}, \qquad \delta v = \frac{\partial \delta\psi}{\partial x}.$$
2. Compute the linearly balanced mass, $\delta^{\mathrm{b}}\tilde{\rho}'$:
$$\delta^{\mathrm{b}}\tilde{\rho}' = \frac{f\delta\psi}{C}.$$
3. Compute the vertically regressed balanced mass:
$$\delta^{\mathrm{b}}_{\mathrm{r}}\tilde{\rho}'_{\mathrm{r}} = \mathbf{R}_{\mathrm{p}}\delta^{\mathrm{b}}\tilde{\rho}'.$$
4. Compute the total mass field:
$$\delta\tilde{\rho}' = \delta^{\mathrm{b}}_{\mathrm{r}}\tilde{\rho}' + \delta^{\mathrm{u}}\tilde{\rho}'.$$
5. Compute the balanced buoyancy, $\delta^{\mathrm{u}}b'$:
$$\delta^{\mathrm{b}}b' = \mathbf{L}^{\mathrm{hb}}\delta\tilde{\rho}'.$$
6. Compute the total buoyancy:
$$\delta b' = \delta^{\mathrm{b}}b' + \delta^{\mathrm{u}}b'.$$
7. Compute the anelastically balanced vertical wind:
$$\delta^{\mathrm{b}}w = \mathbf{L}^{\mathrm{ab}}\delta u.$$
8. Compute the total vertical wind:
$$\delta w = \delta^{\mathrm{b}}w + \delta^{\mathrm{u}}w.$$

9. Compute the tracer:

$$\delta q = \delta q.$$

The hydrostatic balance operator is

$$\mathbf{L}^{\mathrm{hb}}\tilde{\rho}' = C\frac{\partial\tilde{\rho}'}{\partial z},$$

and the anelastic balance operator is

$$\mathbf{L}^{\mathrm{ab}}\delta u = -\frac{1}{\rho_0}\int dz\frac{\partial\rho_0\delta u}{\partial x}.$$

The matrix form of $\mathbf{U}_{\mathrm{p}}$ is

$$\delta\mathbf{x} = \begin{pmatrix} \delta u \\ \delta v \\ \delta\tilde{\rho}' \\ \delta b' \\ \delta w \\ \delta q \end{pmatrix} = \mathbf{U}_{\mathrm{p}}\delta\boldsymbol{\chi} = \begin{pmatrix} 0 & \partial_x & 0 & 0 & 0 & 0 \\ \partial_x & 0 & 0 & 0 & 0 & 0 \\ \mathbf{R}_{\mathrm{p}}\frac{f}{C} & 0 & 1 & 0 & 0 & 0 \\ \mathbf{L}^{\mathrm{hb}}\mathbf{R}_{\mathrm{p}}\frac{f}{C} & 0 & \mathbf{L}^{\mathrm{hb}} & 1 & 0 & 0 \\ 0 & \mathbf{L}^{\mathrm{ab}}\partial_x & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \delta\psi \\ \delta\chi \\ \delta^{\mathrm{u}}\tilde{\rho}' \\ \delta^{\mathrm{u}}b' \\ \delta^{\mathrm{u}}w \\ \delta q \end{pmatrix}.$$

3 Inverse parameter transform

Here is the procedure for the standard inverse transform $\mathbf{U}_{\mathrm{p}}^{-1}$.

1. Compute ψ and χ:

$$\delta\psi = \left(\frac{\partial}{\partial x}\right)^{-2}\frac{\partial\delta v}{\partial x}, \qquad \delta\chi = \left(\frac{\partial}{\partial x}\right)^{-2}\frac{\partial\delta u}{\partial x}.$$

2. Compute the linearly balanced mass, $\delta^{\mathrm{b}}\tilde{\rho}'$:

$$\delta^{\mathrm{b}}\tilde{\rho}' = \frac{f\delta\psi}{C}.$$

3. Compute the vertically regressed balanced mass:

$$\delta_{\mathrm{r}}^{\mathrm{b}}\tilde{\rho}_{\mathrm{r}}' = \mathbf{R}_{\mathrm{p}}\delta^{\mathrm{b}}\tilde{\rho}'.$$

4. Compute the unbalanced mass:

$$\delta^{\mathrm{u}}\tilde{\rho}' = \delta\tilde{\rho}' - \delta_{\mathrm{r}}^{\mathrm{b}}\tilde{\rho}'.$$

5. Compute the balanced buoyancy, $\delta^{\mathrm{u}}b'$:

$$\delta^{\mathrm{b}}b' = \mathbf{L}^{\mathrm{hb}}\delta\tilde{\rho}'.$$

6. Compute the unbalanced buoyancy:

$$\delta^{\mathrm{u}}b' = \delta b' - \delta^{\mathrm{b}}b'.$$

7. Compute the anelastically balanced vertical wind:

$$\delta^{\mathrm{b}}w = \mathbf{L}^{\mathrm{ab}}\delta u.$$

8. Compute the unbalanced vertical wind:

$$\delta^{\mathrm{u}}w = \delta w - \delta^{\mathrm{b}}w.$$

9. Compute the tracer:

$$\delta q = \delta q.$$

The hydrostatic balance operator is

$$\mathbf{L}^{\mathrm{hb}}\tilde{\rho}' = C\frac{\partial \tilde{\rho}'}{\partial z},$$

and the anelastic balance operator is

$$\mathbf{L}^{\mathrm{ab}}\delta u = -\frac{1}{\rho_0}\int dz \frac{\partial \rho_0 \delta u}{\partial x}.$$

The matrix form of $\mathbf{U}_{\mathrm{p}}^{-1}$ is

$$\delta \boldsymbol{\chi} = \begin{pmatrix} \delta\psi \\ \delta\chi \\ \delta^{\mathrm{u}}\tilde{\rho}' \\ \delta^{\mathrm{u}}b' \\ \delta^{\mathrm{u}}w \\ \delta q \end{pmatrix} = \mathbf{U}_{\mathrm{p}}^{-1}\delta\mathbf{x} = \begin{pmatrix} 0 & \partial_x^{-1} & 0 & 0 & 0 & 0 \\ \partial_x^{-1} & 0 & 0 & 0 & 0 & 0 \\ 0 & -\mathbf{R}_{\mathrm{p}}\frac{f}{C}\partial_x^{-1} & 1 & 0 & 0 & 0 \\ 0 & 0 & -\mathbf{L}^{\mathrm{hb}} & 1 & 0 & 0 \\ -\mathbf{L}^{\mathrm{ab}} & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \delta u \\ \delta v \\ \delta\tilde{\rho}' \\ \delta b' \\ \delta w \\ \delta q \end{pmatrix},$$

where

$$\partial_x^{-1} = \left(\frac{\partial}{\partial x}\right)^{-2}\frac{\partial}{\partial x}.$$

4 Adjoint of the anelastic balance constraint

The forward of the anelastic balance operator (vertical motion computed from horizontal winds) is:

$$\delta\Delta(z) = \frac{\partial \rho_0(z)\delta u(z)}{\partial x}$$

$$\delta\rho_w(z) = -\sum_{z'=0}^{z}\delta\Delta(z')d(z')$$

$$\delta w_b(z) = \delta\rho_w(z)/\rho_0(z)$$

$d(z)$ is the level thickness and the active perturbation variables are prefixed with a δ. Let hat variables be the adjoint variables. The adjoint of the last step is

$$\delta\hat{\rho}_w(z) = \delta\hat{w}_b(z)/\rho_0(z).$$

The adjoint of the penultimate step is developed as follows (using the equivalence of the properties of the adjoint with those of the partial derivative):

$$\begin{aligned} \frac{\partial}{\partial\delta\Delta(z')} &= \sum_{z=0}^{\mathrm{top}}\frac{\partial\delta\rho_w(z)}{\partial\delta\Delta(z')}\frac{\partial}{\partial\delta\rho_w(z)} \\ &= \sum_{z=0}^{\mathrm{top}}\left\{\begin{array}{cc} 0 & z' > z \\ -d(z') & \text{otherwise} \end{array}\right\}\frac{\partial}{\partial\delta\rho_w(z)} \\ &= -\sum_{z=z'}^{\mathrm{top}} d(z')\frac{\partial}{\partial\delta\rho_w(z)}. \end{aligned}$$

The adjoint step is then:

$$\delta\hat{\Delta}(z') = -\sum_{z=z'}^{\mathrm{top}} d(z')\delta\hat{\rho}_w(z).$$

5 The vertical regression for balanced pressure

Part of the parameter transform (step 3 of Sect. 2, and step 3 of Sect. 3) is the use of a balanced pressure regression, $\mathbf{R}_{\mathrm{p}}$. This is computed as:

$$\mathbf{R}_{\mathrm{p}} = \mathbf{C}^{\delta\tilde{\rho}',\delta^{\mathrm{b}}\tilde{\rho}'} \left(\mathbf{C}^{\delta^{\mathrm{b}}\tilde{\rho}',\delta^{\mathrm{b}}\tilde{\rho}'}\right)^{-1},$$

where $\mathbf{C}^{\delta^{\mathrm{b}}\tilde{\rho}',\delta^{\mathrm{b}}\tilde{\rho}'}$ is the vertical auto-covariance matrix of the balanced mass (as computed from the linear balance equation), and $\mathbf{C}^{\delta\tilde{\rho}',\delta^{\mathrm{b}}\tilde{\rho}'}$ is the vertical covariance matrix of the total mass with this balanced mass.

6 The Fourier transforms

From the fftpack5 documentation at this link, the definition of the one-dimensional FFT from real $r(i)$ to spectral $s(k)$ space arrays is (assuming that N is even):

$$s(0) = \frac{1}{N}\sum_{i=0}^{N-1} r(i).$$

For $1 \leq k \leq N/2-1$:

$$\begin{aligned} s(2k-1) &= \frac{2}{N}\sum_{i=0}^{N-1} r(i)\cos(2\pi ki/N) \\ s(2k) &= \frac{2}{N}\sum_{i=0}^{N-1} r(i)\sin(2\pi ki/N) \\ s(N-1) &= \frac{1}{N}\sum_{i=0}^{N-1} (-1)^i r(i). \end{aligned}$$

The first value $k=1$ has $2k-1=1$ and $2k=2$; the last value, $k=N/2-1$ has $2k-1=N-3$ and $2k=N-2$.

The above represents an efficient evaluation of the usual FT formula:

$$\hat{s}(k) = \frac{1}{N}\sum_{i=0}^{N-1} r(i)\exp(\iota 2\pi ik/N),$$

where $\iota=\sqrt{-1}$. Given that r is a real function, then $\hat{s}^*(-k)=\hat{s}(k)$, so we don't need to evaluate this for $-N/2 \leq k \leq N/2-1$

Here is how the RFFT1F routine stores the data.

array index, q	0	1	2	3	4	...
wavenumber, k	0	1	1	2	2	
$s(q)$	$\hat{s}(0)$	$2Re\,[\hat{s}(1)]$	$2Im\,[\hat{s}(1)]$	$2Re\,[\hat{s}(2)]$	$2Im\,[\hat{s}(2)]$	
$q(k)$	—	$2k-1$	$2k$	$2k-1$	$2k$	

array index, q	...	$N-3$	$N-2$	$N-1$
wavenumber, k		$N/2-1$	$N/2-1$	$N/2$
$s(q)$		$2Re\,[\hat{s}(N/2-1)]$	$2Im\,[\hat{s}(N/2-1)]$	$\hat{s}(N/2)$
$q(k)$		$2k-1$	$2k$	—

In the convention in my code, I use the q index to go from 1 to N instead of 0 to $N-1$, so the above table becomes the following.

array index, q	1	2	3	4	5	...	$N-2$	$N-1$	N
wavenumber, k	0	1	1	2	2		$N/2-1$	$N/2-1$	$N/2$
$s(q)$	$\hat{s}(0)$	$2Re\,[\hat{s}(1)]$	$2Im\,[\hat{s}(1)]$	$2Re\,[\hat{s}(2)]$	$2Im\,[\hat{s}(2)]$		$2Re\,[\hat{s}(N/2-1)]$	$2Im\,[\hat{s}(N/2-1)]$	$\hat{s}(N/2)$
$q(k)$	—	$2k$	$2k+1$	$2k$	$2k+1$		$2k$	$2k+1$	—

7 What is a function of what in the CVT?

V=Vertical Transform, HT=Horizontal Trabsform, EV=EigenValues, wn=wavenumber, vm=vertical mode.
x and y indicate the axes of a plot, z indicates variable that is repeated over.

Order of transforms	**Nature of VT**		**Horiz EVs**	**Vert modes**	**Vert EVs**
Classic (VT then HT)	symmetric VT	$\Rightarrow$	x: wn, y: height	y: height, z: vm	y: vm
	non-symmetric VT	$\Rightarrow$	x:wn, y:vm	y: height, z: vm	y: vm
Reversed (HT then VT)	symmetric VT	$\Rightarrow$	x: wn, y: height	x: wn, y: height, z: vm	x: wn, y: vm
	non-symmetric VT	$\Rightarrow$	x: wn, y: height	x: wn, y: height, z: vm	x: wn, y: vm

8 Linearization of wind speed

8.1 Horizontal wind speed

The horiztonal wind speed is:

$$s_{\mathrm{h}}^2 = u^2 + v^2.$$

The linearization of this is:

$$\begin{aligned} 2s_{\mathrm{h}}\delta s_{\mathrm{h}} &= 2u\delta u + 2v\delta v \\ \delta s_{\mathrm{h}} &= (u\delta u + v\delta v)/s_{\mathrm{h}} \end{aligned}$$

8.2 Total wind speed

The total wind speed is:

$$s_{\mathrm{t}}^2 = u^2 + v^2 + w^2.$$

The linearization of this is:

$$\begin{aligned} 2s_{\mathrm{t}}\delta s_{\mathrm{t}} &= 2u\delta u + 2v\delta v + 2w\delta w \\ \delta s_t &= (u\delta u + v\delta v + w\delta w)/s_{\mathrm{t}} \end{aligned}$$

9 The gradient of the incremental cost function

9.1 Developing the incremental cost function

Let quantities without a time argument represent that quantity at the start of the window ($t = 0$). The cost function is

$$J(\mathbf{x}) = \frac{1}{2}(\mathbf{x} - \mathbf{x}^{\mathrm{b}})\mathbf{B}^{-1}(\mathbf{x} - \mathbf{x}^{\mathrm{b}}) + \frac{1}{2}\sum_{t=0}^{T} \left[\mathbf{y}(t) - \mathcal{H}_t(\mathcal{M}_{0\to t}(\mathbf{x}))\right]^{\mathrm{T}} \mathbf{R}_t^{-1} \left[\mathbf{y}(t) - \mathcal{H}_t(\mathcal{M}_{0\to t}(\mathbf{x}))\right].$$

Developing this into incremental form:

$$\begin{aligned}
\text{definition: } \mathbf{x}(t) &= \mathcal{M}_{t-1\to t}\left(\mathbf{x}(t-1)\right) \\
\text{definition: } \mathbf{x}(t) &= \mathbf{x}^{\mathrm{R}}(t) + \delta\mathbf{x}(t) \\
\text{definition: } \mathbf{x}^{\mathrm{R}}(t) &= \mathcal{M}_{t-1\to t}\left(\mathbf{x}^{\mathrm{R}}(t-1)\right) \\
\text{putting the above together: } \mathbf{x}^{\mathrm{R}}(t) + \delta\mathbf{x}(t) &= \mathcal{M}_{t-1\to t}\left(\mathbf{x}^{\mathrm{R}}(t-1) + \delta\mathbf{x}(t-1)\right) \\
&\simeq \mathcal{M}_{t-1\to t}\left(\mathbf{x}^{\mathrm{R}}(t-1)\right) + \mathbf{M}_{t-1\to t}\delta\mathbf{x}(t-1) \\
\text{so: } \delta\mathbf{x}(t) &= \mathbf{M}_{t-1\to t}\delta\mathbf{x}(t-1).
\end{aligned}$$

$$\begin{aligned}
\text{definition: } \mathbf{y}^{\mathrm{m}}(t) &= \mathcal{H}_t(\mathbf{x}(t)) \\
\text{definition: } \mathbf{y}^{\mathrm{R}}(t) &= \mathcal{H}_t(\mathbf{x}^{\mathrm{R}}(t)) \\
\text{definition: } \mathbf{y}^{\mathrm{m}}(t) &= \mathbf{y}^{\mathrm{R}}(t) + \delta\mathbf{y}(t) \\
\text{putting the above together: } \mathbf{y}^{\mathrm{m}}(t) &= \mathcal{H}_t\left(\mathbf{x}^{\mathrm{R}}(t) + \delta\mathbf{x}(t)\right) \\
\text{so: } \mathbf{y}^{\mathrm{R}}(t) + \delta\mathbf{y}(t) &\simeq \mathcal{H}_t\left(\mathbf{x}^{\mathrm{R}}(t)\right) + \mathbf{H}_t\delta\mathbf{x}(t) \\
\text{so: } \delta\mathbf{y}(t) &= \mathbf{H}_t\delta\mathbf{x}(t).
\end{aligned}$$

The incremental cost function is then:

$$\begin{aligned}
J(\delta\mathbf{x}) &= \frac{1}{2}\left(\delta\mathbf{x} - \delta\mathbf{x}^{\mathrm{b}}\right)^{\mathrm{T}} \mathbf{B}^{-1}\left(\delta\mathbf{x} - \delta\mathbf{x}^{\mathrm{b}}\right) + \frac{1}{2}\sum_{t=0}^{T}\left[\mathbf{d}(t) - \mathbf{H}_t\mathbf{M}_{0\to t}\delta\mathbf{x}\right]^{\mathrm{T}} \mathbf{R}_t^{-1}\left[\mathbf{d}(t) - \mathbf{H}_t\mathbf{M}_{0\to t}\delta\mathbf{x}\right], \\
&= \frac{1}{2}\left(\delta\mathbf{x} - \delta\mathbf{x}^{\mathrm{b}}\right)^{\mathrm{T}} \mathbf{B}^{-1}\left(\delta\mathbf{x} - \delta\mathbf{x}^{\mathrm{b}}\right) + \frac{1}{2}\sum_{t=0}^{T}\left[\mathbf{H}_t\mathbf{M}_{0\to t}\delta\mathbf{x} - \mathbf{d}(t)\right]^{\mathrm{T}} \mathbf{R}_t^{-1}\left[\mathbf{H}_t\mathbf{M}_{0\to t}\delta\mathbf{x} - \mathbf{d}(t)\right]
\end{aligned}$$

where

$$\begin{aligned}
\delta\mathbf{x}^{\mathrm{b}} &= \mathbf{x}^{\mathrm{b}} - \mathbf{x}^{\mathrm{R}} \\
\mathbf{d}(t) &= \mathbf{y}(t) - \mathcal{H}_t(\mathcal{M}_{0\to t}(\mathbf{x}^{\mathrm{R}})).
\end{aligned}$$

9.2 The derivative of the incremental cost function

The gradient of the incremental cost function (wrt $\delta\mathbf{x}$ and for the reference trajectory $\mathbf{x}^{\mathrm{R}}(t)$) is

$$\nabla_{\delta\mathbf{x}} J(\delta\mathbf{x}) = \mathbf{B}^{-1}\left(\delta\mathbf{x} - \delta\mathbf{x}^{\mathrm{b}}\right) + \sum_{t=0}^{T}\mathbf{M}_{0\to t}^{\mathrm{T}}\mathbf{H}_t^{\mathrm{T}}\mathbf{R}_t^{-1}\left[\mathbf{H}_t\mathbf{M}_{0\to t}\delta\mathbf{x} - \mathbf{d}(t)\right].$$

9.2.1 The adjoint method

The gradient can be made very efficient by considering the following evaluation of the above gradient formula. Note the following definition:

$$\Delta(t) = \mathbf{H}_t^{\mathrm{T}}\mathbf{R}_t^{-1}\left[\mathbf{H}_t\mathbf{M}_{0\to t}\delta\mathbf{x} - \mathbf{d}(t)\right] = \mathbf{H}_t^{\mathrm{T}}\mathbf{R}_t^{-1}\left[\delta\mathbf{y}(t) - \mathbf{d}(t)\right]$$

Time	Forward (non-linear)	Forward (linear)	Adjoint summation									
$t=0$	$\mathbf{x}^{\mathrm{R}}$	$\delta\mathbf{x} = \mathbf{U}\delta\boldsymbol{\chi}$	$\Delta(0)+$									
$t=1$	$\mathbf{x}^{\mathrm{R}}(1) = \mathcal{M}_{0\to1}\left(\mathbf{x}^{\mathrm{R}}\right)$	$\delta\mathbf{x}(1) = \mathbf{M}_{0\to1}\delta\mathbf{x}$	$\mathbf{M}^{\mathrm{T}}_{0\to1}$	$\Delta(1)+$								
$t=2$	$\mathbf{x}^{\mathrm{R}}(2) = \mathcal{M}_{1\to2}\left(\mathbf{x}^{\mathrm{R}}(1)\right)$	$\delta\mathbf{x}(2) = \mathbf{M}_{1\to2}\delta\mathbf{x}(1)$	$\mathbf{M}^{\mathrm{T}}_{0\to1}$	$\mathbf{M}^{\mathrm{T}}_{1\to2}$	$\Delta(2)+$							
$\dots$	$\Downarrow$	$\Downarrow$										
$t-1$	$\mathbf{x}^{\mathrm{R}}(t-1) = \mathcal{M}_{t-2\to t-1}\left(\mathbf{x}^{\mathrm{R}}(t-2)\right)$	$\delta\mathbf{x}(t-1) = \mathbf{M}_{t-2\to t-1}\delta\mathbf{x}(t-2)$	$\mathbf{M}^{\mathrm{T}}_{0\to1}$	$\mathbf{M}^{\mathrm{T}}_{1\to2}$	$\dots$	$\mathbf{M}^{\mathrm{T}}_{t-2\to t-1}$	$\Delta(t-1)+$					
t	$\mathbf{x}^{\mathrm{R}}(t) = \mathcal{M}_{t-1\to t}\left(\mathbf{x}^{\mathrm{R}}(t-1)\right)$	$\delta\mathbf{x}(t) = \mathbf{M}_{t-1\to t}\delta\mathbf{x}(t-1)$	$\mathbf{M}^{\mathrm{T}}_{0\to1}$	$\mathbf{M}^{\mathrm{T}}_{1\to2}$	$\dots$	$\mathbf{M}^{\mathrm{T}}_{t-2\to t-1}$	$\mathbf{M}^{\mathrm{T}}_{t-1\to t}$	$\Delta(t)+$				
$t+1$	$\mathbf{x}^{\mathrm{R}}(t+1) = \mathcal{M}_{t\to t+1}\left(\mathbf{x}^{\mathrm{R}}(t)\right)$	$\delta\mathbf{x}(t+1) = \mathbf{M}_{t\to t+1}\delta\mathbf{x}(t)$	$\mathbf{M}^{\mathrm{T}}_{0\to1}$	$\mathbf{M}^{\mathrm{T}}_{1\to2}$	$\dots$	$\mathbf{M}^{\mathrm{T}}_{t-2\to t-1}$	$\mathbf{M}^{\mathrm{T}}_{t-1\to t}$	$\mathbf{M}^{\mathrm{T}}_{t\to t+1}$	$\Delta(t+1)+$			
$\dots$	$\Downarrow$	$\Downarrow$										
$T-1$	$\mathbf{x}^{\mathrm{R}}(T-1) = \mathcal{M}_{T-2\to T-1}\left(\mathbf{x}^{\mathrm{R}}(T-2)\right)$	$\delta\mathbf{x}(T-1) = \mathbf{M}_{T-2\to T-1}\delta\mathbf{x}(T-2)$	$\mathbf{M}^{\mathrm{T}}_{0\to1}$	$\mathbf{M}^{\mathrm{T}}_{1\to2}$	$\dots$	$\mathbf{M}^{\mathrm{T}}_{t-2\to t-1}$	$\mathbf{M}^{\mathrm{T}}_{t-1\to t}$	$\mathbf{M}^{\mathrm{T}}_{t\to t+1}$	$\dots$	$\mathbf{M}^{\mathrm{T}}_{T-2\to T-1}$	$\Delta(T-1)+$	
T	$\mathbf{x}^{\mathrm{R}}(T) = \mathcal{M}_{T-1\to T}\left(\mathbf{x}^{\mathrm{R}}(T-1)\right)$	$\delta\mathbf{x}(T) = \mathbf{M}_{T-1\to T}\delta\mathbf{x}(T-1)$	$\mathbf{M}^{\mathrm{T}}_{0\to1}$	$\mathbf{M}^{\mathrm{T}}_{1\to2}$	$\dots$	$\mathbf{M}^{\mathrm{T}}_{t-2\to t-1}$	$\mathbf{M}^{\mathrm{T}}_{t-1\to t}$	$\mathbf{M}^{\mathrm{T}}_{t\to t+1}$	$\dots$	$\mathbf{M}^{\mathrm{T}}_{T-2\to T-1}$	$\mathbf{M}^{\mathrm{T}}_{T-1\to T}$	$\Delta(T)$

To compute the gradient evaulate

$$\mathbf{g}(t) = \Delta(t) + \mathbf{M}_{t\to t+1}^{\mathrm{T}}\mathbf{g}(t+1),$$

for $t = T, T-1, \ldots, 1, 0$ where

$$\mathbf{g}(T+1) = 0.$$

These are the equations for 4DVar. For 3D-FGAT make the substitution

$$\mathbf{M}_{t\to t+1} \to \mathbf{I}.$$

For 3DVar, additionally make the substitution

$$\mathcal{M}_{t\to t+1}\left(\mathbf{x}^{\mathrm{R}}(t)\right) \to \mathbf{x}^{\mathrm{R}}(t).$$

9.2.2 Gradient formula in terms of the control vector

Let a model increment be related to a control variable increment as follows

$$\delta\mathbf{x} = \mathbf{U}\delta\boldsymbol{\chi},$$

and specifically for the background increment

$$\delta\mathbf{x}^{\mathrm{b}} = \mathbf{U}\delta\boldsymbol{\chi}^{\mathrm{b}}.$$

The cost function is

$$J(\delta\boldsymbol{\chi}) = \frac{1}{2}\left(\delta\boldsymbol{\chi} - \delta\boldsymbol{\chi}^{\mathrm{b}}\right)^{\mathrm{T}}\left(\delta\boldsymbol{\chi} - \delta\boldsymbol{\chi}^{\mathrm{b}}\right) + \frac{1}{2}\sum_{t=0}^{T}\left[\mathbf{H}_t\mathbf{M}_{0\to t}\mathbf{U}\delta\boldsymbol{\chi} - \mathbf{d}(t)\right]^{\mathrm{T}}\mathbf{R}_t^{-1}\left[\mathbf{H}_t\mathbf{M}_{0\to t}\mathbf{U}\delta\boldsymbol{\chi} - \mathbf{d}(t)\right],$$

and the gradient of the cost function with respect to the control variable $\delta\boldsymbol{\chi}$ is

$$\nabla_{\delta\boldsymbol{\chi}}J(\delta\boldsymbol{\chi}) = \delta\boldsymbol{\chi} - \delta\boldsymbol{\chi}^{\mathrm{b}} + \mathbf{U}^{\mathrm{T}}\sum_{t=0}^{T}\mathbf{M}_{0\to t}^{\mathrm{T}}\mathbf{H}_t^{\mathrm{T}}\mathbf{R}_t^{-1}\left[\mathbf{H}_t\mathbf{M}_{0\to t}\mathbf{U}\delta\boldsymbol{\chi} - \mathbf{d}(t)\right].$$

10 The conjugate gradient method

1. Set the outer loop interation index to zero, $(k) = 0$.
2. Set the inner loop iteration index to zero, $i = 0$.
3. If $k = 0$, set the reference state to the background, $\mathbf{x}_{(k)}^{\mathrm{R}} = \mathbf{x}^{\mathrm{b}}$ (this means that $\delta\mathbf{x}^{\mathrm{b}} = 0$ and $\delta\boldsymbol{\chi}^{\mathrm{b}} = 0$).
4. If $k > 0$, this means that $\delta\mathbf{x}_{(k)}^{\mathrm{b}} = \mathbf{x}^{\mathrm{b}} - \mathbf{x}_{(k)}^{\mathrm{R}}$ and so set $\delta\boldsymbol{\chi}_{(k)}^{\mathrm{b}} = \mathbf{U}^{-1}\delta\mathbf{x}_{(k)}^{\mathrm{b}}$.
5. Set the first estimate of $\mathbf{x}$ as $\mathbf{x}_{(k)}^{\mathrm{R}}$ (this means that $\delta\mathbf{x}_i = 0$ and $\delta\boldsymbol{\chi}_i = 0$).
6. Compute the gradient, $\nabla_{\delta\boldsymbol{\chi}}J(\delta\boldsymbol{\chi}_i, \delta\boldsymbol{\chi}_{(k)}^{\mathrm{b}}) = -\mathbf{r}_i$.
7. Set the initial search direction, $\mathbf{p}_i = \mathbf{r}_i$.
8. Do a line minimization along $\mathbf{p}_i$ to determine $\delta\boldsymbol{\chi}_{i+1}$.
   (a) $\delta\boldsymbol{\chi}_{i+1} = \delta\boldsymbol{\chi}_i + \beta_i\mathbf{p}_i$.
   (b) $J = J(\delta\boldsymbol{\chi}_i) \quad J^+ = J(\delta\boldsymbol{\chi}_i + \mu\mathbf{p}_i) \quad J^- = J(\delta\boldsymbol{\chi}_i - \mu\mathbf{p}_i)$.
   (c) μ is a small number.
   (d) $J(\beta_i) = a\beta_i^2 + b\beta_i + c$.

(e) $dJ(\beta_i)/d\beta_i = 2a\beta_i + b = 0$.

(f) $\beta_i = -b/2a$.

(g) $\beta_i = 0$: $J = c$.

(h) $\beta_i = \mu$: $J^+ = a\mu^2 + b\mu + J$.

(i) $\beta_i = -\mu$: $J^- = a\mu^2 - b\mu + J$.

(j) $J^+ + J^- = 2a\mu^2 + 2J$ and so $a = \left(J^+ + J^- - 2J\right)/2\mu^2$.

(k) $J^+ - J^- = 2b\mu$ and so $b = \left(J^+ - J^-\right)/2\mu$.

(l) Putting a and b into (8f): $\beta_i = (1/2)\left[\left(J^- - J^+\right)/2\mu\right]\left[2\mu^2/\left(J^+ + J^- - 2J\right)\right] = \mu\left(J^- - J^+\right)/\{2\left(J^+ + J^- -\right.$

9. Compute the new gradient, $\nabla_{\delta\boldsymbol{\chi}} J(\delta\boldsymbol{\chi}_{i+1}, \delta\boldsymbol{\chi}^{\mathrm{b}}_{(k)}) = -\mathbf{r}_{i+1}$.

10. Compute the new search direction, $\mathbf{p}_{i+1}$.

    (a) $\mathbf{p}_{i+1} = \mathbf{r}_{i+1} + \alpha_{ii}\mathbf{p}_i$.

    (b) $\alpha_{ii} = \mathbf{r}^{\mathrm{T}}_{i+1}\mathbf{r}_{i+1}/\mathbf{r}^{\mathrm{T}}_i\mathbf{r}_i$.

11. Increment i.

12. Go to (8) until converged.

13. Let $\mathbf{x}^{\mathrm{R}}_{(k+1)} = \mathbf{x}^{\mathrm{R}}_{(k)} + \mathbf{U}\delta\boldsymbol{\chi}_i$.

14. Go to 2 until converged.