Adjoint Model

- Definition of an adjoint operator
  A is B's adjoint if $\vec{x}^T(A\vec{y}) = (B\vec{x})^T\vec{y}$
  for real vectors $\vec{x}, \vec{y}$: $A = B^T$

- Linearized observation operator (H) has its adjoint (H^T)
  Recall OI update equation

$$\delta\vec{x} = BH^T\left((HBH^T+R)^{-1}\vec{d}^{o-b}\right) = BH^T\vec{z}$$

where $\vec{z}$ is the normalized innovation vector

Example: $\vec{x} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$ $\quad \vec{y} = H\vec{x} = \begin{pmatrix} \sigma x_1 \\ \frac{\sigma x_2 + \sigma x_3}{2} \end{pmatrix}$

$$H = \begin{pmatrix} \sigma & 0 & 0 \\ 0 & \frac{\sigma}{2} & \frac{\sigma}{2} \end{pmatrix}$$

$$\begin{pmatrix} \delta x_1 \\ \delta x_2 \\ \delta x_3 \end{pmatrix} = \overline{\vec{\varepsilon}\vec{\varepsilon}^T}\left(H^T\vec{z}\right) = \underbrace{\overline{\vec{\varepsilon}\left(H\vec{\varepsilon}\right)^T}}\vec{z}$$

$$= \begin{pmatrix} \sigma & 0 \\ 0 & \frac{\sigma}{2} \\ 0 & \frac{\sigma}{2} \end{pmatrix}\begin{pmatrix} z_1 \\ z_2 \end{pmatrix} \qquad = \overline{\begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \varepsilon_3 \end{pmatrix}\left(\sigma\varepsilon_1 \quad \frac{\sigma\varepsilon_2 + \sigma\varepsilon_3}{2}\right)}$$

$$= \begin{pmatrix} \sigma z_1 \\ \frac{\sigma}{2} z_2 \\ \frac{\sigma}{2} z_2 \end{pmatrix}$$

As a result: $\delta x_2 = \overline{\varepsilon_2\varepsilon_1}\sigma z_1 + \overline{\varepsilon_2^2}\frac{\sigma}{2}z_2 + \overline{\varepsilon_2\varepsilon_3}\frac{\sigma}{2}z_2$

or $= \sigma\overline{\varepsilon_2\varepsilon_1}\, z_1 + \sigma\,\overline{\varepsilon_2\frac{\varepsilon_1+\varepsilon_2}{2}}\, z_2$

"convert from observation space to state space"

"redistribute innovation to grid points"

→ Don't confuse H^T with the inverse transform $h^{-1}()$!

$$\delta\vec{x} = BH^T\vec{z} = \overline{\varepsilon\varepsilon^T}\, H^T\vec{z} = \overline{\varepsilon(H\varepsilon)^T}\,\vec{z}$$

- on one hand, H^T converts (maps) observed information $\vec{z}$ back to the state space, so that B can be used to make updates $\delta\vec{x}$,
- on the other hand, $\overline{\varepsilon(H\varepsilon)^T}$ can be viewed as the covariance between state (ε) and observation ($H\varepsilon$) spaces.

- The adjoint of TLM is called the "adjoint model" (ADM).
- similar to the functionality of H and H^T, the tangent linear and adjoint models, M and M^T, can be used to map perturbations / increments / sensitivity gradients across time:

$$\delta\vec{x}_t = BM^TH^T\underbrace{\left(HMBM^TH^T+R\right)^{-1}\left(y^o_{t+1} - h(m(x^b_t))\right)}_{=\vec{z}_{t+1}}$$

$$= BM^TH^T\vec{z}_{t+1}$$

$$= \overline{\varepsilon\varepsilon^T}\, M^TH^T\vec{z}_{t+1}$$

$$= \underbrace{\overline{\varepsilon(HM\varepsilon)^T}}_{\text{spatial-temporal error covariance}}\,\vec{z}_{t+1}$$

→ observed information $\vec{z}_{t+1}$ is in observation space and at time $t+1$, $(HM)^T$ maps it back to state space at time t.