PERSONAL NOTES ON DERIVATIVES, INNER PRODUCTS AND ADJOINTS

Ross Bannister, Data Assimilation Research Centre
March 2003

1. CALCULATING THE GRADIENT OF THE OBSERVATIONAL PART OF THE COST FUNCTION

1.1 Matrix notation

$$
\begin{aligned}
J_o &= \frac{1}{2}(\vec{h}(\vec{x}) - \vec{y})^T \mathbf{R}^{-1} (\vec{h}(\vec{x}) - \vec{y}), \\
&= \frac{1}{2}\sum_{ij}(h_i - y_i)(\mathbf{R}^{-1})_{ij}(h_j - y_j), \\
\therefore \frac{\partial J_o}{\partial x_k} &= \frac{1}{2}\sum_{ij}\left(\frac{\partial h_i}{\partial x_k}(\mathbf{R}^{-1})_{ij}(h_j - y_j) + (h_i - y_i)(\mathbf{R}^{-1})_{ij}\frac{\partial h_j}{\partial x_k}\right), \\
&= \sum_{ij}\frac{\partial h_i}{\partial x_k}(\mathbf{R}^{-1})_{ij}(h_j - y_j), \\
&= \sum_{ij}\mathbf{H}_{ik}(\mathbf{R}^{-1})_{ij}(h_j - y_j), \\
\therefore \left(\frac{\partial J_o}{\partial \vec{x}}\right)^T &= \mathbf{H}^T\mathbf{R}^{-1}(\vec{h}(\vec{x}) - \vec{y}).
\end{aligned} \tag{1}
$$

This gives the column of partial derivatives.

1.2 Bra-ket notation

In this notation, we note the following definition,

$$
\begin{aligned}
J_o &= \frac{1}{2}\langle(\vec{h}(\vec{x}) - \vec{y}), \mathbf{R}^{-1}(\vec{h}(\vec{x}) - \vec{y})\rangle, \\
\delta J_o &= \frac{1}{2}\langle\mathbf{H}\delta\vec{x}, \mathbf{R}^{-1}(\vec{h}(\vec{x}) - \vec{y})\rangle + \frac{1}{2}\langle(\vec{h}(\vec{x}) - \vec{y}), \mathbf{R}^{-1}\mathbf{H}\delta\vec{x}\rangle, \\
&= \langle\mathbf{H}\delta\vec{x}, \mathbf{R}^{-1}(\vec{h}(\vec{x}) - \vec{y})\rangle, \\
&= \langle\mathbf{R}^{-1}(\vec{h}(\vec{x}) - \vec{y}), \mathbf{H}\delta\vec{x}\rangle, \\
&= \langle\mathbf{H}^*\mathbf{R}^{-1}(\vec{h}(\vec{x}) - \vec{y}), \delta\vec{x}\rangle.
\end{aligned} \tag{2}
$$

Eq. (3) uses the definition of the adjoint. Compare Eq. (2) to (3),

$$\delta J_o = \langle\nabla_x J_o, \delta\vec{x}\rangle, \tag{3}$$

$$\therefore \nabla_x J_o = \mathbf{H}^*\mathbf{R}^{-1}(\vec{h}(\vec{x}) - \vec{y}). \tag{4}$$

This gives the column of functional (or Gateaux derivatives). These are not generally the same as partial derivatives in Eq. (1). Similarly, the adjoint in Eq. (4) is not generally the same as the transpose in (1).

2. SOME IMPORTANT RELATIONS BETWEEN PARTIAL AND FUNCTIONAL DERIVATIVES, AND TRANSPOSES AND ADJOINTS

2.1 Derivatives

Define a bra-ket inner product as the following (**P** must be symmetric),

$$\langle \vec{u}, \vec{v} \rangle_x = \vec{u}^T \mathbf{P}_x \vec{v}. \tag{5}$$

The following follows from Eq. (3),

$$\delta J_o = \langle \nabla_x J_o, \delta\vec{x} \rangle_x = (\nabla_x J_o)^T \mathbf{P}_x \delta\vec{x}. \tag{6}$$

We also know that δJ_o can be calculated with the partial derivatives,

$$\delta J_o = \frac{\partial J_o}{\partial \vec{x}} \delta\vec{x}, \tag{7}$$

(where the partial derivative vector is a row vector by default). Eqs. (6) and (7) must give the same answer. Comparing yields,

$$\frac{\partial J_o}{\partial \vec{x}} = (\nabla_x J_o)^T \mathbf{P}_x,$$

$$\nabla_x J_o = \mathbf{P}_x^{-1} \left(\frac{\partial J_o}{\partial \vec{x}} \right)^T. \tag{8}$$

This is the relationship between functional and partial derivatives (note that $\mathbf{P}_x$ is symmetric). They are the same if the inner product matrix, $\mathbf{P}_x = \mathbf{I}$.

2.2 Transpose and adjoint (observation operator example)

The definition of the adjoint is that,

$$\langle \delta\vec{y}, \mathbf{H}\delta\vec{x} \rangle_y = \langle \mathbf{H}^* \delta\vec{y}, \delta\vec{x} \rangle_x, \tag{9}$$

where the subscripts on the kets indicate the space (x or y) over which the inner product is performed. Converting Eq. (9) into matrix form means that we have to include the inner product matrices explicitly,

$$\delta\vec{y}^T \mathbf{P}_y \mathbf{H} \delta\vec{x} = (\mathbf{H}^* \delta\vec{y})^T \mathbf{P}_x \delta\vec{x},$$

$$\delta\vec{y}^T \mathbf{P}_y \mathbf{H} \delta\vec{x} = \delta\vec{y}^T (\mathbf{H}^*)^T \mathbf{P}_x \delta\vec{x},$$

$$\mathbf{P}_y \mathbf{H} = (\mathbf{H}^*)^T \mathbf{P}_x,$$

$$\mathbf{H}^T \mathbf{P}_y = \mathbf{P}_x \mathbf{H}^*,$$

$$\mathbf{H}^* = \mathbf{P}_x^{-1} \mathbf{H}^T \mathbf{P}_y. \tag{10}$$

The adjoint is the transpose only in the case of identity inner products.

2.3 Transpose and adjoint (same space operator example)

$$\langle \delta\vec{x}_1, \mathbf{H}\delta\vec{x}_2 \rangle_x = \langle \mathbf{H}^* \delta\vec{x}_1, \delta\vec{x}_2 \rangle_x$$

$$\delta\vec{x}_1^T \mathbf{P}_x \mathbf{H} \delta\vec{x}_2 = (\mathbf{H}^* \delta\vec{x}_1)^T \mathbf{P}_x \delta\vec{x}_2$$

$$\delta\vec{x}_1^T \mathbf{P}_x \mathbf{H} \delta\vec{x}_2 = \delta\vec{x}_1^T (\mathbf{H}^*)^T \mathbf{P}_x \delta\vec{x}_2$$

$$\mathbf{P}_x \mathbf{H} = (\mathbf{H}^*)^T \mathbf{P}_x$$

$$\mathbf{H}^* = \mathbf{P}_x^{-1}\mathbf{H}^T\mathbf{P}_x. \tag{11}$$

3. CHAIN RULES FOR PARTIAL AND FUNCTIONAL DERIVATIVES

3.1 Chain rule for partial derivatives

If,

$$\delta\vec{h} = \mathbf{H}\delta\vec{x}, \tag{12}$$

and the partial derivatives of a scalar with respect to components of $\delta\vec{h}$ are known, then how can the partial derivatives with respect to components of $\delta\vec{x}$ be computed? The following can act on a scalar,

$$\begin{aligned} \frac{\partial}{\partial x_j} &= \sum_i \frac{\partial h_i}{\partial x_j}\frac{\partial}{\partial h_i}, \\ &= \sum_i \mathbf{H}_{ij}\frac{\partial}{\partial h_i}, \\ \therefore \left(\frac{\partial}{\partial \vec{x}}\right)^T &= \mathbf{H}^T\left(\frac{\partial}{\partial h}\right)^T. \end{aligned} \tag{13}$$

This is the chain rule for partial derivatives.

3.2 Chain rule for functional derivatives

With reference to Eq. (8) applied to the x and h derivatives, Eq. (13) is transformed to,

$$\mathbf{P}_x\nabla_x = \mathbf{H}^T\mathbf{P}_y\nabla_y.$$

Using now Eq. (10) to rewrite the transpose operator, this becomes,

$$\begin{aligned} \mathbf{P}_x\nabla_x &= \mathbf{P}_x\mathbf{H}^*\mathbf{P}_y^{-1}\mathbf{P}_y\nabla_y, \\ \nabla_x &= \mathbf{H}^*\nabla_y, \end{aligned} \tag{14}$$

which is the chain rule for functional derivatives, of a similar structure to Eq. (13) for partial derivatives.