Complex Gradient and Hessian
(A. van den Bos, IEE Proc. Vis. Imag Proc., 141(6) 1994)

Let:

$w \in \mathbb{R}^{2N}$, $w = (x_1, y_1, x_2, y_2, \cdots, x_N, y_N)^T$

consider: $f: \mathbb{R}^{2N} \to \mathbb{R}$ "smooth"

$$f(w) \approx f(0) + \underbrace{\frac{\partial f}{\partial w^T} w}_{①} + \frac{1}{2} w^T \frac{\partial^2 f}{\partial w \partial w^T} w$$

$\frac{\partial f}{\partial w} \triangleq$ "gradient"

$\frac{\partial^2 f}{\partial w \partial w^T} \triangleq$ "Hessian"

① linear term

② quadratic term

Next define $z_n \triangleq x_n + j y_n$ $\quad z_n^* = x_n - j y_n$

$$\begin{pmatrix} z_n \\ z_n^* \end{pmatrix} = \begin{pmatrix} 1 & j \\ 1 & -j \end{pmatrix} \begin{pmatrix} x_n \\ y_n \end{pmatrix}$$

$$J \triangleq \begin{pmatrix} 1 & j \\ 1 & -j \end{pmatrix} \qquad J^{-1} = \frac{1}{2} \begin{pmatrix} 1 & 1 \\ -j & j \end{pmatrix} = \frac{1}{2} J^H$$

thus, we see that z_n and z_n^* are independent variables.

$$\therefore \begin{pmatrix} x_n \\ y_n \end{pmatrix} = \frac{1}{2} J^H \begin{pmatrix} z_n \\ z_n^* \end{pmatrix}$$

Now define $\quad v \triangleq (z_1, z_1^*, z_2, z_2^* \cdots z_N, z_N^*)$

$$\Rightarrow \boxed{v = A\, w} \quad \text{where} \quad A \triangleq \operatorname{diag}(J, \ldots J)$$

$$A^{-1} = \operatorname{diag}(J^{-1}, \ldots, J^{-1}) \Rightarrow A^{-1} = \frac{1}{2} A^H$$

$$\therefore \boxed{w = \frac{1}{2} A^H v} \quad \text{but } w \text{ is real, so}$$

$$\boxed{w = \frac{1}{2} A^T v^*}$$

Now consider, $\dfrac{\partial f}{\partial w^T} w$, the linear term.

$$\frac{\partial f}{\partial w^T} w = \underbrace{\frac{1}{2} \frac{\partial f}{\partial w^T} A^H}_{\dfrac{\partial f}{\partial v^T}} v \qquad \text{linear term in } f(v)$$

$$\boxed{\frac{\partial f}{\partial v^T} = \frac{1}{2} \frac{\partial f}{\partial w^T} A^H} \Rightarrow \boxed{\frac{\partial f}{\partial v} = \frac{1}{2} A^* \frac{\partial f}{\partial w}}$$

complex gradient of a real function w.r.t. v.

$$A^{-1} = \frac{1}{2} A^H$$

$$\frac{\partial f}{\partial v^T} = \frac{\partial f}{\partial w^T} A^{-1}$$

$$\frac{\partial f}{\partial v^T} A = \frac{\partial f}{\partial w^T} \Rightarrow \boxed{A^T \frac{\partial f}{\partial v} = \frac{\partial f}{\partial w}}$$

real gradient

Properties

From $\frac{\partial f}{\partial v} = \frac{1}{2} A^* \frac{\partial f}{\partial w}$ we get

$$\left.\begin{aligned} \frac{\partial f}{\partial z_n} &= \frac{1}{2}\left(\frac{\partial f}{\partial x_n} - j\frac{\partial f}{\partial y_n}\right) \\ \frac{\partial f}{\partial z_n^*} &= \frac{1}{2}\left(\frac{\partial f}{\partial x_n} + j\frac{\partial f}{\partial y_n}\right) \end{aligned}\right\} \left(\frac{\partial f}{\partial z_n}\right)^* = \frac{\partial f}{\partial z_n^*}$$

$$\therefore \quad \frac{\partial f}{\partial v^*} = \left(\frac{\partial f}{\partial v}\right)^*$$

but $\left(\frac{\partial f}{\partial w}\right)^* = \frac{\partial f}{\partial w}$ because f is real and so is w

$$= \left(A^T \frac{\partial f}{\partial v}\right)^* = A^H \frac{\partial f}{\partial v^*}$$

or $\frac{\partial f}{\partial w^T} = \frac{\partial f}{\partial v^H} A^*$

Thus we have:

$$\boxed{\frac{\partial f}{\partial w} = 0 \iff \frac{\partial f}{\partial v} = 0 \iff \frac{\partial f}{\partial v^*} = 0}$$

Complex Hessian

$\frac{1}{2} w^T H w$ → quadratic term

$$H = \frac{\partial^2 f}{\partial w \partial w^T}$$

but $w = \frac{1}{2} A^H v = \frac{1}{2} A^T v^*$

∴ quadratic term in v:

$$\frac{1}{8} v^H A H A^H v = \frac{1}{2} v^T \frac{\partial^2 f}{\partial v \partial v^T} v$$

$$= \frac{1}{2} v^H \frac{\partial^2 f}{\partial v^* \partial v^T} v$$

∴ $$\boxed{G \triangleq \frac{\partial^2 f}{\partial v^* \partial v^T} = \frac{1}{4} A H A^H}$$

"Complex Hessian"

but $A^{-1} = \frac{1}{2} A^H$ ∴ $$\boxed{H = \frac{\partial^2 f}{\partial w \partial w^T} = A^H G A}$$

Notes on the adjoint method of determining the derivative of a functional with respect to a parameter.

Consider the operator equation:

$$A \underline{x} = \underline{s}$$

where A is a matrix, $\underline{x}$ is the unknown vector and $\underline{s}$ is the source vector.

Formally, $\underline{x} = A^{-1} \underline{s}$

(But we may not want to calculate A^{-1}!)

Suppose that what we want is to calculate the rate of change of a functional with respect to a parameter h.

Consider the functional: $\phi = \underline{d}^T \underline{x}$

$$\frac{\partial \phi}{\partial h} = \underline{d}^T \frac{\partial \underline{x}}{\partial h}$$

Now sp that both A and $\underline{s}$ depend on some parameter h.

I.e., $A(h) \underline{x} = \underline{s}(h)$

taking the derivative :

$$A \frac{\partial \underline{x}}{\partial h} + \frac{\partial A}{\partial h} \underline{x} = \frac{\partial \underline{s}}{\partial h}$$

$$A \frac{\partial \underline{x}}{\partial h} = -\left(\frac{\partial A}{\partial h} \underline{x} - \frac{\partial \underline{s}}{\partial h} \right)$$

$$\frac{\partial \underline{x}}{\partial h} = -A^{-1} \left(\frac{\partial A}{\partial h} \underline{x} - \frac{\partial \underline{s}}{\partial h} \right)$$

$$\frac{\partial \phi}{\partial h} = -\underline{d}^T A^{-1} \left(\frac{\partial A}{\partial h} \underline{x} - \frac{\partial \underline{s}}{\partial h} \right)$$

$\frac{\partial \underline{s}}{\partial h}, \frac{\partial A}{\partial h}$ – we should be able to find analytically.

Thus, the multiplication $\underline{d}^T A^{-1} \left(\frac{\partial A}{\partial h} \underline{x} \right)$ could be a problem, in terms of computational expense.

So, what can be done is to define the "adjoint" solution to the "adjoint" problem as $A^T \underline{x}^a = -\underline{d} \Rightarrow (\underline{x}^a)^T = -\underline{d}^T A^{-1}$

$$\therefore \boxed{\frac{\partial \phi}{\partial h} = (\underline{x}^a)^T \left(\frac{\partial A}{\partial h} \underline{x} - \frac{\partial \underline{s}}{\partial h} \right)}$$