Vector and Matrix Derivatives

Introduction

This lecture supplement is not meant to be an exhaustive compilation all vector and matrix derivatives. Instead, it presents (with proofs) several commonly used results in digital signal processing and estimation theory, for both real– and complex–valued vectors and matrices. Students should be able to use the examples of the proofs presented in this supplement as a starting point for many other derivatives of scalars and vectors with respect to vectors and matrices. For a more comprehensive list of matrix identities, we refer to [1, 2].

Real Valued Vectors and Matrices

Definitions:

- Column vectors with elements $\in \mathbb{R}$
  - $\mathbf{y} = \left[y_1, y_2, \ldots, y_N\right]^T$, $\mathbf{x} = \left[x_1, x_2, \ldots, x_M\right]^T$, $\mathbf{z} = \left[z_1, z_2, \ldots, z_K\right]^T$, and $\mathbf{a} = \left[a_1, a_2, \ldots, a_M\right]^T$.
- General $(N \times M)$ matrix $\mathbf{A}$ and $(M \times M)$ square matrices $\mathbf{B}$ and $\mathbf{C}$ where

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1M} \\ a_{21} & a_{22} & \cdots & a_{2M} \\ \vdots & \vdots & \ddots & \vdots \\ a_{N1} & a_{N2} & \cdots & a_{NM} \end{bmatrix} \qquad \mathbf{C} = \begin{bmatrix} c_{11} & c_{12} & \cdots & c_{1M} \\ c_{21} & c_{22} & \cdots & c_{2M} \\ \vdots & \vdots & \ddots & \vdots \\ c_{M1} & c_{M2} & \cdots & c_{MM} \end{bmatrix} \qquad \mathbf{B} = \begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1M} \\ b_{21} & b_{22} & \cdots & b_{2M} \\ \vdots & \vdots & \ddots & \vdots \\ b_{M1} & b_{M2} & \cdots & b_{MM} \end{bmatrix}$$

- The partial derivatives of a scalar α with respect to a vector $\mathbf{x}$ and matrix $\mathbf{A}$ are defined as

$$\frac{\partial \alpha}{\partial \mathbf{x}} = \begin{bmatrix} \frac{\partial \alpha}{\partial x_1} \\ \frac{\partial \alpha}{\partial x_2} \\ \vdots \\ \frac{\partial \alpha}{\partial x_M} \end{bmatrix}_{M \times 1} \qquad \frac{\partial \alpha}{\partial \mathbf{A}} = \begin{bmatrix} \frac{\partial \alpha}{\partial a_{11}} & \frac{\partial \alpha}{\partial a_{12}} & \cdots & \frac{\partial \alpha}{\partial a_{1M}} \\ \frac{\partial \alpha}{\partial a_{21}} & \frac{\partial \alpha}{\partial a_{22}} & \cdots & \frac{\partial \alpha}{\partial a_{2M}} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial \alpha}{\partial a_{N1}} & \frac{\partial \alpha}{\partial a_{N2}} & \cdots & \frac{\partial \alpha}{\partial a_{NM}} \end{bmatrix}_{N \times M} \tag{1}$$

- The partial derivative of a vector $\mathbf{y}$ with respect to vector $\mathbf{x}$ is[1].

$$\frac{\partial \mathbf{y}}{\partial \mathbf{x}} \triangleq \frac{\partial \mathbf{y}^T}{\partial \mathbf{x}} = \begin{bmatrix} \frac{\partial y_1}{\partial x_1} & \frac{\partial y_2}{\partial x_1} & \cdots & \frac{\partial y_N}{\partial x_1} \\ \frac{\partial y_1}{\partial x_2} & \frac{\partial y_2}{\partial x_2} & \cdots & \frac{\partial y_N}{\partial x_2} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial y_1}{\partial x_M} & \frac{\partial y_2}{\partial x_M} & \cdots & \frac{\partial y_N}{\partial x_M} \end{bmatrix}_{M \times N} \tag{2}$$

Summary

$$\text{Linear Form:} \quad \frac{\partial}{\partial \mathbf{x}} \left\{\mathbf{x}^T \mathbf{a}\right\} = \frac{\partial}{\partial \mathbf{x}} \left\{\mathbf{a}^T \mathbf{x}\right\} = \mathbf{x} \tag{3}$$

$$\text{Quadratic Form:} \quad \frac{\partial}{\partial \mathbf{x}} \left\{\mathbf{x}^T \mathbf{C} \mathbf{x}\right\} = \left(\mathbf{C} + \mathbf{C}^T\right) \mathbf{x} \overset{\star}{=} 2\mathbf{C}\mathbf{x}, \quad {}^\star\text{for symmetric } \mathbf{C} \tag{4}$$

$$\text{Vector Form:} \quad \mathbf{y} = \mathbf{A}\mathbf{x}, \quad \frac{\partial \mathbf{y}^T}{\partial \mathbf{x}} = \mathbf{A}^T \tag{5}$$

Chain Rule

$$\text{Linear Form:} \quad \frac{\partial}{\partial \mathbf{z}} \left\{\mathbf{a}^T \mathbf{x}\right\} = \frac{\partial \mathbf{x}^T}{\partial \mathbf{z}} \mathbf{a} + \frac{\partial \mathbf{a}^T}{\partial \mathbf{z}} \mathbf{x} \tag{6}$$

$$\text{Quadratic Form:} \quad \frac{\partial}{\partial \mathbf{z}} \left\{\mathbf{x}^T \mathbf{C} \mathbf{x}\right\} = \frac{\partial \mathbf{x}^T}{\partial \mathbf{z}} \left(\mathbf{C} + \mathbf{C}^T\right) \mathbf{x} \overset{\star}{=} 2 \frac{\partial \mathbf{x}^T}{\partial \mathbf{z}} \mathbf{C}\mathbf{x}, \quad {}^\star\text{for symmetric } \mathbf{C} \tag{7}$$

$$\text{Vector Form:} \quad \mathbf{y} = \mathbf{A}\mathbf{x}, \quad \frac{\partial \mathbf{y}^T}{\partial \mathbf{z}} = \frac{\partial \mathbf{x}^T}{\partial \mathbf{z}} \mathbf{A}^T \tag{8}$$

[1] Be aware that some texts define $\frac{\partial \mathbf{y}}{\partial \mathbf{x}} = \frac{\partial \mathbf{y}}{\partial \mathbf{x}^T}$, which would alter the results of the derivatives.

Matrix Derivatives

$$\text{Linear Form:} \quad \frac{\partial}{\partial \mathbf{B}} \left\{ \mathsf{Tr}\left(\mathbf{BC}\right) \right\} = \mathbf{C}^T \tag{9}$$

$$\text{Quadratic Form:} \quad \frac{\partial}{\partial \mathbf{A}} \left\{ \mathsf{Tr}\left(\mathbf{ACA}^T\right) \right\} = \mathbf{A}\left(\mathbf{C}^T + \mathbf{C}\right) \overset{\star}{=} 2\mathbf{AC}, \quad {}^{\star}\text{for symmetric } \mathbf{C} \tag{10}$$

Complex Valued Vectors and Matrices

Definitions:

- Complex-valued column vectors $\mathbf{y} \in \mathbb{C}^{N\times 1}$, $\mathbf{x} \in \mathbb{C}^{M\times 1}$, $\mathbf{z} \in \mathbb{C}^{K\times 1}$ and $\mathbf{a} \in \mathbb{C}^{M\times 1}$.
- Matrices $\mathbf{A} \in \mathbb{C}^{N\times M}$, $\mathbf{B} \in \mathbb{C}^{M\times M}$ and $\mathbf{C} \in \mathbb{C}^{M\times M}$.
- The partial derivative of a scalar α with respect a vector $\mathbf{x}^*$ and matrix $\mathbf{A}^*$ is

$$\frac{\partial \alpha}{\partial \mathbf{x}^*} = \begin{bmatrix} \frac{\partial \alpha}{\partial x_1^*} \\ \frac{\partial \alpha}{\partial x_2^*} \\ \vdots \\ \frac{\partial \alpha}{\partial x_M^*} \end{bmatrix}_{M\times 1} \qquad \frac{\partial \alpha}{\partial \mathbf{A}^*} = \begin{bmatrix} \frac{\partial \alpha}{\partial a_{11}} & \frac{\partial \alpha}{\partial a_{12}^*} & \cdots & \frac{\partial \alpha}{\partial a_{1M}^*} \\ \frac{\partial \alpha}{\partial a_{21}^*} & \frac{\partial \alpha}{\partial a_{22}^*} & \cdots & \frac{\partial \alpha}{\partial a_{2M}^*} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial \alpha}{\partial a_{N1}^*} & \frac{\partial \alpha}{\partial a_{N2}^*} & \cdots & \frac{\partial \alpha}{\partial a_{NM}^*} \end{bmatrix}_{N\times M} \tag{11}$$

- Unlike the real-valued case, the partial derivative of a complex valued vector can be defined in two ways

$$\frac{\partial \mathbf{y}^T}{\partial \mathbf{x}^*} = \begin{bmatrix} \frac{\partial y_1}{\partial x_1^*} & \frac{\partial y_2}{\partial x_1^*} & \cdots & \frac{\partial y_N}{\partial x_1^*} \\ \frac{\partial y_1}{\partial x_2^*} & \frac{\partial y_2}{\partial x_2^*} & \cdots & \frac{\partial y_N}{\partial x_2^*} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial y_1}{\partial x_M^*} & \frac{\partial y_2}{\partial x_M^*} & \cdots & \frac{\partial y_N}{\partial x_M^*} \end{bmatrix}_{M\times N} \qquad \frac{\partial \mathbf{y}^H}{\partial \mathbf{x}^*} = \begin{bmatrix} \frac{\partial y_1^*}{\partial x_1^*} & \frac{\partial y_2^*}{\partial x_1^*} & \cdots & \frac{\partial y_N^*}{\partial x_1^*} \\ \frac{\partial y_1^*}{\partial x_2^*} & \frac{\partial y_2^*}{\partial x_2^*} & \cdots & \frac{\partial y_N^*}{\partial x_2^*} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial y_1^*}{\partial x_M^*} & \frac{\partial y_2^*}{\partial x_M^*} & \cdots & \frac{\partial y_N^*}{\partial x_M^*} \end{bmatrix}_{M\times N} \tag{12}$$

Conjugate Gradient

The complex valued derivatives used in this text are not the conventional complex derivatives used in the standard mathematics and engineering complex variables courses. This is because standard complex derivative operators are only defined for holomorphic functions that obey the Cauchy-Riemann conditions [3]. Since many useful functions in engineering are non-holormopic (e.g. squared error cost function), we resort to the partial derivatives operators from the so-called $\mathbb{CR}$–calculus framework. For more information, we refer the reader to the tutorial in [3].

Summary

$$\text{Linear Form:} \quad \frac{\partial}{\partial \mathbf{x}^*}\left\{\mathbf{x}^T\mathbf{a}\right\} = \mathbf{0} \tag{13}$$

$$\text{Linear Form:} \quad \frac{\partial}{\partial \mathbf{x}^*}\left\{\mathbf{x}^H\mathbf{a}\right\} = \mathbf{a} \tag{14}$$

$$\text{Quadratic Form:} \quad \frac{\partial}{\partial \mathbf{x}^*}\left\{\mathbf{x}^H\mathbf{C}\mathbf{x}\right\} = \mathbf{C}\mathbf{x} \tag{15}$$

$$\text{Quadratic Form:} \quad \frac{\partial}{\partial \mathbf{x}^*}\left\{\mathbf{x}^T\mathbf{C}\mathbf{x}^*\right\} = \mathbf{C}^T\mathbf{x} \tag{16}$$

$$\text{Vector Form:} \quad \mathbf{y} = \mathbf{A}\mathbf{x}, \quad \frac{\partial \mathbf{y}^H}{\partial \mathbf{x}^*} = \mathbf{A}^H \tag{17}$$

Chain Rule

$$\text{Linear Form:} \quad \frac{\partial}{\partial \mathbf{z}^*}\left\{\mathbf{x}^H\mathbf{a}\right\} = \frac{\partial \mathbf{x}^H}{\partial \mathbf{z}^*}\mathbf{a} + \frac{\partial \mathbf{a}^T}{\partial \mathbf{z}^*}\mathbf{x}^* \tag{18}$$

$$\text{Quadratic Form:} \quad \frac{\partial}{\partial \mathbf{z}^*}\left\{\mathbf{x}^H\mathbf{C}\mathbf{x}\right\} = \frac{\partial \mathbf{x}^H}{\partial \mathbf{z}^*}\mathbf{C}\mathbf{x} + \frac{\partial \mathbf{x}^T}{\partial \mathbf{z}^*}\mathbf{C}^T\mathbf{x}^* \tag{19}$$

$$\text{Vector Form:} \quad \mathbf{y} = \mathbf{A}\mathbf{x}, \quad \frac{\partial \mathbf{y}^H}{\partial \mathbf{z}^*} = \frac{\partial \mathbf{x}^H}{\partial \mathbf{z}^*}\mathbf{A}^H, \quad \frac{\partial \mathbf{y}^T}{\partial \mathbf{z}^*} = \frac{\partial \mathbf{x}^T}{\partial \mathbf{z}^*}\mathbf{A}^T \tag{20}$$

Matrix Derivatives

$$\text{Linear Form:} \quad \frac{\partial}{\partial \mathbf{B}^*}\left\{\mathsf{Tr}\left(\mathbf{B}^*\mathbf{C}\right)\right\} = \mathbf{C}^T \tag{21}$$

$$\text{Quadratic Form:} \quad \frac{\partial}{\partial \mathbf{A}^*}\left\{\mathsf{Tr}\left(\mathbf{A}\mathbf{C}\mathbf{A}^H\right)\right\} = \mathbf{A}\mathbf{C} \tag{22}$$

Proofs – Real Case

We will prove the results for the chain rules in (6) and (8) and show that the proofs for (3) – (5) and (7) are special cases of (6) and (8).

Linear Form

Proof of (6): Consider the inner product of the vectors $\mathbf{a}$ and $\mathbf{x}$, that is the sum of the products of the elements of the two vectors:

$$\mathbf{a}^T\mathbf{x} = \sum_{\ell=1}^{M} a_\ell x_\ell \tag{23}$$

The k–th element of the partial derivative vector is $\left[\frac{\partial}{\partial \mathbf{z}}\left\{\mathbf{a}^T\mathbf{x}\right\}\right]_k = \frac{\partial}{\partial z_k}\left\{\mathbf{a}^T\mathbf{x}\right\}$ where

$$\frac{\partial}{\partial z_k}\left\{\mathbf{a}^T\mathbf{x}\right\} = \sum_{\ell=1}^{M} \frac{\partial x_\ell}{\partial z_k} a_\ell + \sum_{\ell=1}^{M} \frac{\partial a_\ell}{\partial z_k} x_\ell \tag{24}$$

$$= \sum_{\ell=1}^{M} \left[\frac{\partial \mathbf{x}^T}{\partial \mathbf{z}}\right]_{k\ell} [\mathbf{a}]_\ell + \sum_{\ell=1}^{M} \left[\frac{\partial \mathbf{a}^T}{\partial \mathbf{z}}\right]_{k\ell} [\mathbf{x}]_\ell \tag{25}$$

$$\implies \frac{\partial}{\partial \mathbf{z}}\left\{\mathbf{a}^T\mathbf{x}\right\} = \frac{\partial \mathbf{x}^T}{\partial \mathbf{z}}\mathbf{a} + \frac{\partial \mathbf{a}^T}{\partial \mathbf{z}}\mathbf{x} \tag{26}$$

Proof of (3): a special case of (26) where $\mathbf{z} = \mathbf{x}$ therefore

$$\frac{\partial}{\partial \mathbf{x}}\left\{\mathbf{a}^T\mathbf{x}\right\} = \mathbf{a} \tag{27}$$

since $\frac{\partial \mathbf{x}^T}{\partial \mathbf{x}} = \mathbf{I}$ and the vector $\mathbf{a}$ is independent of $\mathbf{x}$, this gives $\frac{\partial \mathbf{a}^T}{\partial \mathbf{x}} = \mathbf{0}$.

Vector Form

Proof of (8): For the equation $\mathbf{y} = \mathbf{A}\mathbf{x}$, the i–th element of vector $\mathbf{y}$ is the inner product between the i–th row of the matrix $\mathbf{A}$ and the vector $\mathbf{x}$, such that

$$y_i = \sum_{\ell=1}^{M} a_{i\ell} x_\ell \tag{28}$$

The ki–th element (k–th row, i–th column) of the partial derivative matrix $\frac{\partial \mathbf{y}^T}{\partial \mathbf{z}}$ is then given by

$$\frac{\partial y_i}{\partial z_k} = \sum_{\ell=1}^{M} a_{i\ell} \frac{\partial x_\ell}{\partial z_k} \tag{29}$$

$$= \sum_{\ell=1}^{M} \left[\frac{\partial \mathbf{x}^T}{\partial \mathbf{z}}\right]_{k\ell} [\mathbf{A}]_{i\ell} \tag{30}$$

$$\left[\frac{\partial \mathbf{y}^T}{\partial \mathbf{z}}\right]_{ki} = \sum_{\ell=1}^{M} \left[\frac{\partial \mathbf{x}^T}{\partial \mathbf{z}}\right]_{k\ell} \left[\mathbf{A}^T\right]_{\ell i} \tag{31}$$

$$\implies \frac{\partial \mathbf{y}^T}{\partial \mathbf{z}} = \frac{\partial \mathbf{x}^T}{\partial \mathbf{z}}\mathbf{A}^T \tag{32}$$

Proof of (5): The result in (5) is a special case of (32) where $\mathbf{z} = \mathbf{x}$ giving

$$\frac{\partial \mathbf{y}^T}{\partial \mathbf{x}} = \mathbf{A}^T \tag{33}$$

Quadratic Form

Proof of (7): The the partial derivative $\frac{\partial}{\partial \mathbf{z}}\left\{\mathbf{x}^T\mathbf{C}\mathbf{x}\right\}$ can be derived using the results in (26) by defining $\mathbf{a} \triangleq \mathbf{C}^T\mathbf{x}$

$$\frac{\partial}{\partial \mathbf{z}}\left\{\mathbf{x}^T\mathbf{C}\mathbf{x}\right\} = \frac{\partial}{\partial \mathbf{z}}\left\{\mathbf{a}^T\mathbf{x}\right\} = \frac{\partial \mathbf{x}^T}{\partial \mathbf{z}}\mathbf{a} + \frac{\partial \mathbf{a}^T}{\partial \mathbf{z}}\mathbf{x} \tag{34}$$

Using the chain rule result for the vector form in (32), we have

$$\frac{\partial \mathbf{a}^T}{\partial \mathbf{z}} = \frac{\partial \mathbf{x}^T}{\partial \mathbf{z}}\mathbf{C}$$

Therefore

$$\frac{\partial}{\partial \mathbf{z}}\left\{\mathbf{x}^T\mathbf{C}\mathbf{x}\right\} = \frac{\partial \mathbf{x}^T}{\partial \mathbf{z}}\mathbf{C}^T\mathbf{x} + \frac{\partial \mathbf{x}^T}{\partial \mathbf{z}}\mathbf{C}\mathbf{x} = \frac{\partial \mathbf{x}^T}{\partial \mathbf{z}}\left(\mathbf{C}^T + \mathbf{C}\right)\mathbf{x} \tag{35}$$

If the matrix $\mathbf{C}$ is symmetric i.e. $\mathbf{C}^T = \mathbf{C}$, then

$$\frac{\partial}{\partial \mathbf{z}}\left\{\mathbf{x}^T\mathbf{C}\mathbf{x}\right\} = 2\frac{\partial \mathbf{x}^T}{\partial \mathbf{z}}\mathbf{C}\mathbf{x} \tag{36}$$

Proof of (4): From , when $\mathbf{z} = \mathbf{x}$, we obtain the well-known result

$$\frac{\partial}{\partial \mathbf{x}}\left\{\mathbf{x}^T\mathbf{C}\mathbf{x}\right\} = \left(\mathbf{C}^T + \mathbf{C}\right)\mathbf{x} = 2\mathbf{C}\mathbf{x}, \quad \text{for symmetric } \mathbf{C} \tag{37}$$

Matrix Derivatives – Linear Form

Proof of (9): The trace of a matrix, denoted by the operator $\mathsf{Tr}\,(\cdot)$, is the sum of all its diagonal elements. $\mathsf{Tr}\,(\mathbf{BC})$ can be expressed as

$$\mathsf{Tr}\,(\mathbf{BC}) = \sum_{\ell}[\mathbf{BC}]_{\ell\ell} = \sum_{\ell}\sum_{m}[\mathbf{B}]_{\ell m}[\mathbf{C}]_{m\ell} \tag{38}$$

The jk–th element of the partial derivative matrix $\frac{\partial}{\partial \mathbf{B}}\{\mathsf{Tr}\,(\mathbf{BC})\}$ is given by

$$\frac{\partial}{\partial b_{jk}}\{\mathsf{Tr}\,(\mathbf{BC})\} = \sum_{\ell}\sum_{m}\frac{\partial b_{\ell m}}{\partial b_{jk}}[\mathbf{C}]_{m\ell} \tag{39}$$

Assuming all the elements of $\mathbf{B}$ are independent of each other, we have

$$\frac{\partial b_{\ell m}}{\partial b_{jk}} = \begin{cases} 1, & \ell = j, m = k \\ 0, & \text{otherwise} \end{cases}$$

Therefore

$$\left[\frac{\partial \alpha}{\partial \mathbf{A}}\right]_{jk} = [\mathbf{C}]_{kj} \tag{40}$$

$$\Longrightarrow \frac{\partial \alpha}{\partial \mathbf{A}} = \mathbf{C}^T \tag{41}$$

Matrix Derivatives – Quadratic Form

Proof of (10): Using the property $\mathsf{Tr}\,(\mathbf{BC}) = \mathsf{Tr}\,(\mathbf{CB})$, we have

$$\mathsf{Tr}\left(\mathbf{A}\mathbf{C}\mathbf{A}^T\right) = \mathsf{Tr}\left(\mathbf{A}^T\mathbf{A}\mathbf{C}\right) \tag{42}$$

$$= \sum_{\ell}\left[\mathbf{A}^T\mathbf{A}\mathbf{C}\right]_{\ell\ell} = \sum_{\ell}\sum_{m}\left[\mathbf{A}^T\mathbf{A}\right]_{\ell m}[\mathbf{C}]_{m\ell} \tag{43}$$

The jk–th element of the partial derivative matrix $\frac{\partial}{\partial \mathbf{A}}\left\{\mathsf{Tr}\left(\mathbf{A}\mathbf{C}\mathbf{A}^T\right)\right\}$ is

$$\frac{\partial}{\partial a_{jk}}\left\{\mathsf{Tr}\left(\mathbf{A}\mathbf{C}\mathbf{A}^T\right)\right\} = \; = \sum_{\ell}\sum_{m}\frac{\partial[\mathbf{A}^T\mathbf{A}]_{\ell m}}{\partial a_{jk}}[\mathbf{C}]_{m\ell} \tag{44}$$

The matrix derivative

$$\frac{\partial[\mathbf{A}^T\mathbf{A}]_{\ell m}}{\partial a_{jk}} = \frac{\partial}{\partial a_{jk}} \sum_n [\mathbf{A}^T]_{\ell n}[\mathbf{A}]_{nm} \tag{45}$$

$$= \frac{\partial}{\partial a_{jk}} \sum_n [\mathbf{A}]_{n\ell}[\mathbf{A}]_{nm} \tag{46}$$

$$= \sum_n \frac{\partial a_{n\ell} a_{nm}}{\partial a_{jk}} \tag{47}$$

$$= \frac{\partial a_{j\ell} a_{jm}}{\partial a_{jk}} \tag{48}$$

$$= \begin{cases} 2a_{jk}, & \ell = m = k \\ a_{jm}, & \ell = k, \;\; m \neq k \\ a_{j\ell}, & m = k, \;\; \ell \neq k \end{cases} \tag{49}$$

Substituting (49) into (44) gives

$$\sum_\ell \sum_m \frac{\partial[\mathbf{A}^T\mathbf{A}]_{\ell m}}{\partial a_{jk}} [\mathbf{C}]_{m\ell} = 2a_{jk} [\mathbf{C}]_{kk} + \sum_{\ell \neq k} a_{j\ell} [\mathbf{C}]_{k\ell} + \sum_{m \neq k} a_{jm} [\mathbf{C}]_{mk} \tag{50}$$

$$= \sum_\ell a_{j\ell} [\mathbf{C}]_{k\ell} + \sum_m a_{jm} [\mathbf{C}]_{mk} \tag{51}$$

$$= \sum_\ell [\mathbf{A}]_{j\ell} \left[\mathbf{C}^T\right]_{\ell k} + \sum_m [\mathbf{A}]_{jm} [\mathbf{C}]_{mk} \tag{52}$$

$$\tag{53}$$

Therefore, we have the result

$$\frac{\partial}{\partial a_{jk}} \left\{ \mathsf{Tr}\left(\mathbf{A}\mathbf{C}\mathbf{A}^T\right) \right\} = [\mathbf{A}\mathbf{C}^T]_{jk} + [\mathbf{A}\mathbf{C}]_{jk} \tag{54}$$

$$\implies \frac{\partial}{\partial \mathbf{A}} \left\{ \mathsf{Tr}\left(\mathbf{A}\mathbf{C}\mathbf{A}^T\right) \right\} = \mathbf{A}\left(\mathbf{C}^T + \mathbf{C}\right) = 2\mathbf{A}\mathbf{C}, \quad \text{for symmetric } \mathbf{C} \tag{55}$$

Proofs – Complex Case

Similar to the methodology used in the proofs of the real case, we will first prove the general result for the chain rules, then deduce the other results.

Linear Form

Proof of (18): Firstly, the inner product $\mathbf{x}^H\mathbf{a}$ is expressed as

$$\mathbf{x}^H\mathbf{a} = \sum_{\ell=1}^{M} x_\ell^* a_\ell \tag{56}$$

The k–th element of the vector of partial derivatives $\frac{\partial}{\partial \mathbf{z}^*}\left\{\mathbf{x}^H\mathbf{a}\right\}$ is

$$\frac{\partial}{\partial z_k^*}\left\{\mathbf{x}^H\mathbf{a}\right\} = \sum_{\ell=1}^{M} \frac{\partial x_\ell^*}{\partial z_k^*} a_\ell + \sum_{\ell=1}^{M} \frac{\partial a_\ell}{\partial z_k^*} x_\ell^* \tag{57}$$

$$= \sum_{\ell=1}^{M} \left[\frac{\partial \mathbf{x}^H}{\partial \mathbf{z}^*}\right]_{k\ell} [\mathbf{a}]_\ell + \sum_{\ell=1}^{M} \left[\frac{\partial \mathbf{a}^T}{\partial \mathbf{z}^*}\right]_{k\ell} [\mathbf{x}]_\ell^* \tag{58}$$

$$\implies \frac{\partial}{\partial \mathbf{z}}\left\{\mathbf{x}^H\mathbf{a}\right\} = \frac{\partial \mathbf{x}^H}{\partial \mathbf{z}^*}\mathbf{a} + \frac{\partial \mathbf{a}^T}{\partial \mathbf{z}^*}\mathbf{x}^* \tag{59}$$

Proof of (13): By definition, the partial derivative of $\mathbf{x}^T\mathbf{a}$ is taken with respect to $\mathbf{x}^*$ by keeping $\mathbf{x}$ constant, it is clear that

$$\frac{\partial}{\partial \mathbf{x}^*}\left\{\mathbf{x}^T\mathbf{a}\right\} = \mathbf{0} \tag{60}$$

Proof of (14): If $\mathbf{z}^* = \mathbf{x}^*$, we can see that (59) is

$$\frac{\partial}{\partial \mathbf{x}^*}\left\{\mathbf{x}^H\mathbf{a}\right\} = \mathbf{a} \tag{61}$$

Vector Form

Proof of (20): For the equation $\mathbf{y} = \mathbf{A}\mathbf{x}$, the i–th element of vector $\mathbf{y}$ is the inner product between the i–th row of $\mathbf{A}$ and the vector $\mathbf{x}$, such that

$$y_i = \sum_{\ell=1}^{M} a_{i\ell} x_\ell$$

The ki–th element (k–th row, i–th column) of the matrix of partial derivatives $\frac{\partial}{\partial \mathbf{z}^*}\left\{\mathbf{y}^H\right\}$

$$\frac{\partial y_i^*}{\partial z_k^*} = \sum_{\ell=1}^{M} a_{i\ell}^* \frac{\partial x_\ell^*}{\partial z_k^*} \tag{62}$$

$$= \sum_{\ell=1}^{M} \left[\frac{\partial \mathbf{x}^H}{\partial \mathbf{z}^*}\right]_{k\ell} [\mathbf{A}^*]_{i\ell} \tag{63}$$

$$\left[\frac{\partial \mathbf{y}^H}{\partial \mathbf{z}^*}\right]_{ki} = \sum_{\ell=1}^{M} \left[\frac{\partial \mathbf{x}^H}{\partial \mathbf{z}^*}\right]_{k\ell} \left[\mathbf{A}^H\right]_{\ell i} \tag{64}$$

$$\implies \frac{\partial \mathbf{y}^H}{\partial \mathbf{z}^*} = \frac{\partial \mathbf{x}^H}{\partial \mathbf{z}^*}\mathbf{A}^H \tag{65}$$

Proof of (17): For the special case $\mathbf{z}^* = \mathbf{x}^*$, we have

$$\frac{\partial \mathbf{y}^H}{\partial \mathbf{z}^*} = \mathbf{A}^H \tag{66}$$

Quadratic Form

In the complex-valued case, quadratic forms can take the form of either $\mathbf{x}^H\mathbf{C}\mathbf{x}$ and $\mathbf{x}^T\mathbf{C}\mathbf{x}^*$. (Note: We will not consider the the partial derivatives of $\mathbf{x}^H\mathbf{C}\mathbf{x}^*$ or $\mathbf{x}^T\mathbf{C}\mathbf{x}$, but it will be a useful exercise to the reader to carry out these derivations.)

Proof of (19): For the case of $\frac{\partial}{\partial \mathbf{z}^*}\left\{\mathbf{x}^H\mathbf{C}\mathbf{x}\right\}$ we will first define $\mathbf{a} \triangleq \mathbf{C}\mathbf{x}$ and use the result in (59) to yield

$$\frac{\partial}{\partial \mathbf{z}^*}\left\{\mathbf{x}^H\mathbf{C}\mathbf{x}\right\} = \frac{\partial \mathbf{x}^H}{\partial \mathbf{z}^*}\mathbf{a} + \frac{\partial \mathbf{a}^T}{\partial \mathbf{z}^*}\mathbf{x}^*$$

Then using (65) we find an expression for $\frac{\partial \mathbf{a}^T}{\partial \mathbf{z}^*}$

$$\frac{\partial}{\partial \mathbf{z}^*}\left\{\mathbf{a}^T\right\} = \frac{\partial \mathbf{x}^T}{\partial \mathbf{z}^*}\mathbf{C}^T$$

Therefore

$$\frac{\partial}{\partial \mathbf{z}^*}\left\{\mathbf{x}^H\mathbf{C}\mathbf{x}\right\} = \frac{\partial \mathbf{x}^H}{\partial \mathbf{z}^*}\mathbf{C}\mathbf{x} + \frac{\partial \mathbf{x}^T}{\partial \mathbf{z}^*}\mathbf{C}^T\mathbf{x}^* \tag{67}$$

Proof of (15): For the special case of $\mathbf{z}^* = \mathbf{x}^*$, we have

$$\frac{\partial}{\partial \mathbf{x}^*}\left\{\mathbf{x}^H\mathbf{C}\mathbf{x}\right\} = \mathbf{C}\mathbf{x} \tag{68}$$

Proof of (16): The partial differential of the quadratic form $\mathbf{x}^T\mathbf{C}\mathbf{x}^*$, is

$$\frac{\partial}{\partial \mathbf{z}^*}\left\{\mathbf{x}^T\mathbf{C}\mathbf{x}^*\right\} = \frac{\partial \mathbf{x}^H}{\partial \mathbf{z}^*}\mathbf{C}^T\mathbf{x} + \frac{\partial \mathbf{x}^T}{\partial \mathbf{z}^*}\mathbf{C}\mathbf{x}^* \tag{69}$$

$$\frac{\partial}{\partial \mathbf{x}^*}\left\{\mathbf{x}^T\mathbf{C}\mathbf{x}^*\right\} = \mathbf{C}^T\mathbf{x} \tag{70}$$

Matrix Derivatives – Linear Form

Proof of (21): First, $\mathsf{Tr}\left(\mathbf{B}^*\mathbf{C}\right)$ is expressed as

$$\mathsf{Tr}\left(\mathbf{B}^*\mathbf{C}\right) = \sum_{\ell} \left[\mathbf{B}^*\mathbf{C}\right]_{\ell\ell} \tag{71}$$

$$= \sum_{\ell}\sum_{m} \left[\mathbf{B}^*\right]_{\ell m}\left[\mathbf{C}\right]_{m\ell} \tag{72}$$

The jk–th element of the matrix of partial derivatives with respect to $\mathbf{B}^*$ is

$$\left[\frac{\partial}{\partial \mathbf{B}^*}\left\{\mathsf{Tr}\left(\mathbf{B}^*\mathbf{C}\right)\right\}\right]_{jk} = \sum_{\ell}\sum_{m} \frac{\partial b^*_{\ell m}}{\partial b^*_{jk}}\left[\mathbf{C}\right]_{m\ell} \tag{73}$$

Since $\frac{\partial b^*_{\ell m}}{\partial b^*_{jk}} = 1$ only for $\ell = j$ and $m = k$, we have

$$\left[\frac{\partial}{\partial \mathbf{B}^*}\left\{\mathsf{Tr}\left(\mathbf{B}^*\mathbf{C}\right)\right\}\right]_{jk} = \left[\mathbf{C}\right]_{kj} \tag{74}$$

$$\Longrightarrow \frac{\partial}{\partial \mathbf{B}^*}\left\{\mathsf{Tr}\left(\mathbf{B}^*\mathbf{C}\right)\right\} = \mathbf{C}^T \tag{75}$$

Matrix Derivatives – Quadratic Form

Proof of (22): Using the identity $\mathsf{Tr}\left(\mathbf{C}\mathbf{B}\right) = \mathsf{Tr}\left(\mathbf{B}\mathbf{C}\right)$, the quadratic form of the trace expression is

$$\mathsf{Tr}\left(\mathbf{A}\mathbf{C}\mathbf{A}^H\right) = \mathsf{Tr}\left(\mathbf{A}^H\mathbf{A}\mathbf{C}\right) = \sum_{\ell}\sum_{m} [\mathbf{A}^H\mathbf{A}]_{\ell m}[\mathbf{C}]_{m\ell} \tag{76}$$

The jk–th element of the matrix of partial derivatives with respect to $\mathbf{A}^*$ is

$$\left[\frac{\partial}{\partial \mathbf{A}^*}\left\{\mathsf{Tr}\left(\mathbf{A}^H\mathbf{A}\mathbf{C}\right)\right\}\right]_{jk} = \sum_{\ell}\sum_{m} \frac{\partial [\mathbf{A}^H\mathbf{A}]_{\ell m}}{\partial a^*_{jk}}[\mathbf{C}]_{m\ell} \tag{77}$$

$$\tag{78}$$

The partial derivative of $\mathbf{A}^H\mathbf{A}$ with respect to the individual matrix elements a_{jk}^* is

$$\frac{\partial[\mathbf{A}^H\mathbf{A}]_{\ell m}}{\partial a_{jk}^*} = \sum_n \frac{\partial a_{n\ell}^* a_{nm}}{\partial a_{jk}^*} \tag{79}$$

$$= \begin{cases} a_{jm}, & n=j, \quad \ell=k \\ 0, & \text{otherwise} \end{cases} \tag{80}$$

As a consequence

$$\left[\frac{\partial}{\partial \mathbf{A}^*}\left\{\mathsf{Tr}\left(\mathbf{A}^H\mathbf{A}\mathbf{C}\right)\right\}\right]_{jk} = \sum_m a_{jm}[\mathbf{C}]_{mk} = [\mathbf{A}\mathbf{C}]_{jk} \tag{81}$$

$$\implies \frac{\partial}{\partial \mathbf{A}^*}\left\{\mathsf{Tr}\left(\mathbf{A}^H\mathbf{A}\mathbf{C}\right)\right\} = \mathbf{A}\mathbf{C} \tag{82}$$

References

[1] K. B. Petersen, M. S. Pedersen *et al.*, "The matrix cookbook," *Technical University of Denmark*, vol. 450, pp. 7–15, 2008.

[2] A. Hjorungnes and D. Gesbert, "Complex-valued matrix differentiation: Techniques and key results," *IEEE Transactions on Signal Processing*, vol. 55, no. 6, pp. 2740–2746, 2007.

[3] K. Kreutz-Delgado, "The complex gradient operator and the CR-calculus," *arXiv preprint arXiv:0906.4835*, 2009.