The Guide for Matrix Calculus

Rubem Vasconcelos Pacelli
rubem.engenharia@gmail.com

Department of Teleinformatics Engineering, Federal University of Ceará.
Fortaleza, Ceará, Brazil.

Version: March 22, 2024

Contents

1 Introduction

Since my Master's degree, I've been struggling with matrix differentiation as I could not find good references that cover it nicely. The bibliographies I found at that time were books from Economics [1], but they use an ~~weird~~ unfamiliar notation.

After delving a lot, I finally found a good reference from Professor Randal's class note [2]. However, to my surprise, when I tried to apply those matrix differentiation propositions, I got "wrong" answers! The truth is that *Matrix Calculus notation is severely fragmented and there is no consensus among the researchers over which notation to follow.* Fortunately, there are only two major ways to represent a derivative of a vector [3]. If you do not select the author's representation, you will end up with the transposed result. The first representation is called Jacobian formulation or numerator layout, while the second one is called Hessian formulation or denominator layout. Nevertheless, even for the same layout, some conventions need to be stated.

Due to the lack of references and the need to have one, I decided to make this guide. The goal here is twofold: introduce the theoretical aspects of Matrix Calculus (first part) and solve partial derivatives for the most common Matrix Calculus expressions you may come across (second part). The second part is nothing more than a manual solution for some expressions, which might be useful to the reader as a quick reminder of how to solve them. The step-by-step is given. Most problems will be defined in real space, but some will be extended to complex space. Furthermore, **I will adopt only the demodulator layout** since it matches the notation commonly used by authors in the field of Signal Processing, Machine Learning, and Optimization Theory. On the other hand, the first part might be valuable for those who are trying to understand theoretical concepts, conventions, notations, identities, and applications[1] of Matrix Calculus.

1.1 Curated reference list

The following list shows some references you can rely on besides this guide when it comes to Matrix Calculus (in decreasing order of importance):

- Both layouts:
  - Wikipedia [4]: Despite not being a good practice to cite Wikipedia, it contains surprisingly comprehensive content regarding Matrix Calculus.
- Denominator layout:
  - Searle, Shayle R., and Andre I. Khuri. Matrix algebra useful for statistics [5]: A classical book with the first edition in 1987. It treats Matrix Calculus in chapter 9 of the second edition.
  - Hjørungnes, Are, and David Gesbert. Complex-valued matrix differentiation: Techniques and key results [6]. An advanced book that uses Wirtinger calculus to differentiate complex-valued matrices. You might prefer his article [7] which summarizes the key results.
  - You can find good references in the following book appendices: Dattoro [8], appendix D; Bishop [9], appendix C; Simon Haykin [10], appendix B (Wirtinger calculus).
  - The Matrix Cookbook [11]: Good reference to see Matrix Calculus expression results rather than understanding how they were solved.
  - Lecture notes in Introduction to Machine Learning, from Carnegie Mellon University [3]: A short lecture but with a clear explanation of the difference between numerator and denominator layout.
- Numerator Layout:
  - Dhrymes, Phoebus J and Dhrymes, Phoebus J. Mathematics for econometrics. The classical book for Matrix Calculus. It was originally issued in 1978, and since then it became a main reference for the numerator layout.

[1] Especially to Optimization Theory.

– MP Deisenroth, AA Faisal, CS Ong. Mathematics for Machine Learning [12]: A more recent book that addresses Matrix Calculus for Machine Learning.

– Old and New Matrix Algebra Useful for Statistics, Thomas Minka [13]: Another good guide for matrix algebra hosted on the personal site of the author. It is very well-referenced and focuses more on the concepts of Matrix Calculus than solving the derivatives.

– Matrix Calculus You Need For Deep Learning [14]: Preprint article focused on its application on Machine Learning. It offers the best tradeoff between comprehension and brevity as it is just 33 pages long. It uses the numerator layout, though.

– Professor Randal's class note [2]: It has the same purpose as this guide, that is, to solve common matrix differentiation, but it is done using the numerator layout. Such a layout is rarely used in the references herein cited.

Differentiation solutions that were collected from other sources will be referenced, while solutions that I derived by myself will not have any reference. Obviously, this guide may have errors (I hope not). If you find it, feel free to reach out through email or simply make a pull request on my Github.

Part I
Theory

2 Notation and nomeclatures

Let

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix} \in \mathbb{R}^{m \times n} \tag{1}$$

be a complex matrix with dimension equal to $m \times n$, where $a_{ij} \in \mathbb{R}$ is its element in the position (i, j). Similarly, a complex vector is defined by

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \in \mathbb{R}^n, \tag{2}$$

which may also be denoted as an n-tuple, $(x_1, x_2, \dots, x_n)$, when more convenient.

For scalars, italic Roman $(a, b, c, \dots)$ represents constants (known values), while italic Greek $(\alpha, \beta, \gamma, \dots)$ represents variables (unknown values). For vectors, Roman bold lowercase letters will be adopted, where the initial letters $(\mathbf{a}, \mathbf{b}, \mathbf{c}, \dots)$ represent constants (known values) and the final letters $(\mathbf{x}, \mathbf{y}, \mathbf{w}, \dots)$ represent variables (unknown values). The same thing goes for matrices, but it is denoted as uppercase letters instead of lowercase. The letter $\mathbf{z}$ (or $\mathbf{Z}$, for matrices) will only be used when the vector (or matrix) is complex-valued. In any other case, the vector or matrix will be real-valued. Finally, the operators $\cdot^{\top}$, $\cdot^{\mathsf{H}}$, $\cdot^{*}$ $\mathrm{tr}(\cdot)$, $\mathrm{adj}(\cdot)$, and $|\cdot|$ denote, respectively, the transpose, the hermitian, the conjugate, the trace, the adjoint, and the determinant (or absolute value when the operand is a scalar).

2.1 Jacobian formulation (numerator layout)

In the Jacobian formulation (also called numerator layout), the derivative matrix is written laying out the numerator in its shape, while the denominator has its shape transposed.

2.1.1 Vector-vector, scalar-vector, and vector-scalar derivatives

Consider two vectors $\mathbf{x} \in \mathbb{R}^n$ and $\mathbf{y} \in \mathbb{R}^m$. The partial derivative of each element in $\mathbf{y}$ by each element in $\mathbf{x}$ is represented as

$$\left[\frac{\partial \mathbf{y}}{\partial \mathbf{x}}\right]_{\text{Num}} = \begin{bmatrix} \frac{\partial y_1}{\partial \mathbf{x}} \\ \frac{\partial y_2}{\partial \mathbf{x}} \\ \vdots \\ \frac{\partial y_m}{\partial \mathbf{x}} \end{bmatrix} = \begin{bmatrix} \frac{\partial y_1}{\partial x_1} & \frac{\partial y_1}{\partial x_2} & \cdots & \frac{\partial y_1}{\partial x_n} \\ \frac{\partial y_2}{\partial x_1} & \frac{\partial y_2}{\partial x_2} & \cdots & \frac{\partial y_2}{\partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial y_m}{\partial x_1} & \frac{\partial y_m}{\partial x_2} & \cdots & \frac{\partial y_m}{\partial x_n} \end{bmatrix} \in \mathbb{R}^{m \times n}. \tag{3}$$

We can infer what is the shape of $\frac{\partial y}{\partial \mathbf{x}}$ and $\frac{\partial \mathbf{y}}{\partial x}$ by changing the respective vector sizes in (3).

2.1.2 Matrix-scalar derivative (tangent matrix)

The partial derivative $\frac{\partial \mathbf{Y}}{\partial x}$ (usually called tangent matrix) is defined for the numerator layout as

$$\left[\frac{\partial \mathbf{Y}}{\partial x}\right]_{\text{Num}} = \begin{bmatrix} \frac{\partial y_{11}}{\partial x} & \frac{\partial y_{12}}{\partial x} & \cdots & \frac{\partial y_{1n}}{\partial x} \\ \frac{\partial y_{21}}{\partial x} & \frac{\partial y_{22}}{\partial x} & \cdots & \frac{\partial y_{2n}}{\partial x} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial y_{m1}}{\partial x} & \frac{\partial y_{m2}}{\partial x} & \cdots & \frac{\partial y_{mn}}{\partial x} \end{bmatrix} \in \mathbb{R}^{m \times n}, \tag{4}$$

where $\mathbf{Y} \in \mathbb{R}^{m \times n}$.

2.1.3 Scalar-matrix derivative (gradient matrix)

The partial derivative of $\frac{\partial y}{\partial \mathbf{X}}$ (usually called gradient matrix) is given by

$$\left[\frac{\partial y}{\partial \mathbf{X}}\right]_{\text{Num}} = \begin{bmatrix} \frac{\partial y}{\partial x_{11}} & \frac{\partial y}{\partial x_{21}} & \cdots & \frac{\partial y}{\partial x_{m1}} \\ \frac{\partial y}{\partial x_{12}} & \frac{\partial y}{\partial x_{22}} & \cdots & \frac{\partial y}{\partial x_{m2}} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial y}{\partial x_{1n}} & \frac{\partial y}{\partial x_{2n}} & \cdots & \frac{\partial y}{\partial x_{mn}} \end{bmatrix} \in \mathbb{R}^{n \times m}, \tag{5}$$

where $\mathbf{X} \in \mathbb{R}^{m \times n}$.

2.1.4 Row vector-scalar and scalar-row vector derivatives

From these definitions, we can infer two nonobvious equalities that are rather useful when handling matrix differentiations. If we consider a special case of the gradient matrix (5), when $m = 1 \therefore \mathbf{X} =$

$\mathbf{x}^\top \in \mathbb{R}^{1\times n}$, where $\mathbf{x} \in \mathbb{R}^n$, we have that

$$\left[\frac{\partial \alpha}{\partial \mathbf{x}^\top}\right]_{\text{Num}} = \begin{bmatrix} \frac{\partial \alpha}{\partial x_1} \\ \frac{\partial \alpha}{\partial x_2} \\ \vdots \\ \frac{\partial \alpha}{\partial x_n} \end{bmatrix} \in \mathbb{R}^n. \tag{6}$$

However, by comparing with (3), one can state that

$$\left[\frac{\partial \alpha}{\partial \mathbf{x}}\right]_{\text{Num}} = \begin{bmatrix} \frac{\partial \alpha}{\partial x_1} & \frac{\partial \alpha}{\partial x_2} & \cdots & \frac{\partial \alpha}{\partial x_n} \end{bmatrix} \in \mathbb{R}^{1\times n}, \tag{7}$$

Therefore,

$$\boxed{\left[\frac{\partial \alpha}{\partial \mathbf{x}^\top}\right]_{\text{Num}} = \left[\frac{\partial \alpha}{\partial \mathbf{x}}\right]_{\text{Num}}^\top.} \tag{8}$$

Similarly, from (4), when $m = 1 \therefore \mathbf{Y} = \mathbf{y}^\top \in \mathbb{R}^{1\times n}$, where $\mathbf{y} \in \mathbb{R}^n$, we have that

$$\left[\frac{\partial \mathbf{y}^\top}{\partial \alpha}\right]_{\text{Num}} = \begin{bmatrix} \frac{\partial y_1}{\partial \alpha} & \frac{\partial y_2}{\partial \alpha} & \cdots & \frac{\partial y_n}{\partial \alpha} \end{bmatrix} \in \mathbb{R}^{1\times n}. \tag{9}$$

However, from (3), we also have that

$$\left[\frac{\partial \mathbf{y}}{\partial \alpha}\right]_{\text{Num}} = \begin{bmatrix} \frac{\partial y_1}{\partial \alpha} \\ \frac{\partial y_2}{\partial \alpha} \\ \cdots \\ \frac{\partial y_n}{\partial \alpha} \end{bmatrix} \in \mathbb{R}^n. \tag{10}$$

Therefore,

$$\boxed{\left[\frac{\partial \mathbf{y}^\top}{\partial \alpha}\right]_{\text{Num}} = \left[\frac{\partial \mathbf{y}}{\partial \alpha}\right]_{\text{Num}}^\top.} \tag{11}$$

2.1.5 Jacobian matrix for the numerator layout

In the numerator layout, the notation for the Jacobian matrix is given by

$$\mathbf{J} = \left[\frac{\partial f(\mathbf{x})}{\partial \mathbf{x}}\right]_{\text{Num}}. \tag{12}$$

where $f : \mathbb{R}^n \to \mathbb{R}^m$ is a vector function. By calling (3), it is clear that the Jacobian takes the form

$$\mathbf{J} = \begin{bmatrix} \frac{\partial f_1}{\partial \mathbf{x}} \\ \frac{\partial f_2}{\partial \mathbf{x}} \\ \vdots \\ \frac{\partial f_m}{\partial \mathbf{x}} \end{bmatrix} = \begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \frac{\partial f_2}{\partial x_1} & \frac{\partial f_2}{\partial x_2} & \cdots & \frac{\partial f_2}{\partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial f_m}{\partial x_1} & \frac{\partial f_m}{\partial x_2} & \cdots & \frac{\partial f_m}{\partial x_n} \end{bmatrix} \in \mathbb{R}^{m \times n}, \tag{13}$$

where $f = (f_1, f_2, \ldots, f_m)$, being $f_i : \mathbb{R}^n \to \mathbb{R}$, for $1 \leq i \leq m$. The numerator layout is also called the Jacobian formulation due to the fact that it is represented without the need for the transpose operator.

2.1.6 Hessian matrix for the numerator layout

The Matrix Calculus notation for the Hessian matrix in the numerator layout is given by

$$\mathbf{H} = \left[\left(\frac{\partial^2 f(\mathbf{x})}{\partial \mathbf{x} \partial \mathbf{x}^\top} \right)^\top \right]_{\text{Num}} = \begin{bmatrix} \frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_1 \partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_1 \partial x_n} \\ \frac{\partial^2 f}{\partial x_2 \partial x_1} & \frac{\partial^2 f}{\partial x_2^2} & \cdots & \frac{\partial^2 f}{\partial x_2 \partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 f}{\partial x_n \partial x_1} & \frac{\partial^2 f}{\partial x_n \partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_n^2} \end{bmatrix} \in \mathbb{R}^{n \times n}, \tag{14}$$

where $f : \mathbb{R}^n \to \mathbb{R}$ and $\mathbf{x} \in \mathbb{R}^n$.

The proof of (14) can be seen as follows

$$\left(\frac{\partial^2 f(\mathbf{x})}{\partial \mathbf{x} \partial \mathbf{x}^\top}\right)^\top = \left(\frac{\partial}{\partial \mathbf{x}}\left(\frac{\partial f(\mathbf{x})}{\partial \mathbf{x}^\top}\right)\right)^\top \tag{15}$$

$$= \left(\frac{\partial}{\partial \mathbf{x}}\begin{bmatrix} \frac{\partial f(\mathbf{x})}{\partial x_1} \\ \frac{\partial f(\mathbf{x})}{\partial x_2} \\ \vdots \\ \frac{\partial f(\mathbf{x})}{\partial x_n} \end{bmatrix}\right)^\top \tag{16}$$

$$= \begin{bmatrix} \frac{\partial^2 f(\mathbf{x})}{\partial x_1^2} & \frac{\partial^2 f(\mathbf{x})}{\partial x_2 \partial x_1} & \cdots & \frac{\partial^2 f(\mathbf{x})}{\partial x_n \partial x_1} \\ \frac{\partial^2 f(\mathbf{x})}{\partial x_1 \partial x_2} & \frac{\partial^2 f(\mathbf{x})}{\partial x_2^2} & \cdots & \frac{\partial^2 f(\mathbf{x})}{\partial x_2 \partial x_1} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 f(\mathbf{x})}{\partial x_1 \partial x_n} & \frac{\partial^2 f(\mathbf{x})}{\partial x_2 \partial x_n} & \cdots & \frac{\partial^2 f(\mathbf{x})}{\partial x_n^2} \end{bmatrix}^\top \tag{17}$$

$$= \begin{bmatrix} \frac{\partial^2 f(\mathbf{x})}{\partial x_1^2} & \frac{\partial^2 f(\mathbf{x})}{\partial x_1 \partial x_2} & \cdots & \frac{\partial^2 f(\mathbf{x})}{\partial x_1 \partial x_n} \\ \frac{\partial^2 f(\mathbf{x})}{\partial x_2 \partial x_1} & \frac{\partial^2 f(\mathbf{x})}{\partial x_2^2} & \cdots & \frac{\partial^2 f(\mathbf{x})}{\partial x_1 \partial x_2} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 f(\mathbf{x})}{\partial x_n \partial x_1} & \frac{\partial^2 f(\mathbf{x})}{\partial x_n \partial x_2} & \cdots & \frac{\partial^2 f(\mathbf{x})}{\partial x_n^2} \end{bmatrix} \tag{18}$$

$$= \mathbf{H} \tag{19}$$

Although it has no difference for real-valued derivates as Clairaut's theorem ensures that [15] $\frac{\partial^2 f}{\partial x_a \partial x_b} = \frac{\partial^2 f}{\partial x_b \partial x_a}$, we shall obey the differentiation order as defined in (14) for the sake of consistency.

2.2 Hessian formulation (denominator layout)

In the Hessian formulation (also called denominator layout), the derivative matrix is written laying out the denominator in its shape, while the numerator has its shape transposed.

2.2.1 Vector-vector, scalar-vector, and vector-scalar derivatives

Consider two vectors $\mathbf{x} \in \mathbb{R}^n$ and $\mathbf{y} \in \mathbb{R}^m$. The partial derivative of each element in $\mathbf{y}$ by each element in $\mathbf{x}$ is represented as

$$\left[\frac{\partial \mathbf{y}}{\partial \mathbf{x}}\right]_{\text{Den}} = \begin{bmatrix} \frac{\partial y_1}{\partial \mathbf{x}} & \frac{\partial y_2}{\partial \mathbf{x}} & \cdots & \frac{\partial y_3}{\partial \mathbf{x}} \end{bmatrix} = \begin{bmatrix} \frac{\partial y_1}{\partial x_1} & \frac{\partial y_2}{\partial x_1} & \cdots & \frac{\partial y_m}{\partial x_1} \\ \frac{\partial y_1}{\partial x_2} & \frac{\partial y_2}{\partial x_2} & \cdots & \frac{\partial y_m}{\partial x_2} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial y_1}{\partial x_n} & \frac{\partial y_2}{\partial x_n} & \cdots & \frac{\partial y_m}{\partial x_n} \end{bmatrix} \in \mathbb{R}^{n \times m}. \tag{20}$$

We can infer what is the shape of $\frac{\partial y}{\partial \mathbf{x}}$ and $\frac{\partial \mathbf{y}}{\partial x}$ by changing the respective vector sizes in (20).

2.2.2 Matrix-scalar derivative (tangent matrix)

The partial derivative $\dfrac{\partial \mathbf{Y}}{\partial x}$ (usually called the tangent matrix) is defined for the denominator layout as

$$\left[\frac{\partial \mathbf{Y}}{\partial x}\right]_{\text{Den}} = \begin{bmatrix} \dfrac{\partial y_{11}}{\partial x} & \dfrac{\partial y_{21}}{\partial x} & \cdots & \dfrac{\partial y_{m1}}{\partial x} \\ \dfrac{\partial y_{12}}{\partial x} & \dfrac{\partial y_{22}}{\partial x} & \cdots & \dfrac{\partial y_{m2}}{\partial x} \\ \vdots & \vdots & \ddots & \vdots \\ \dfrac{\partial y_{1n}}{\partial x} & \dfrac{\partial y_{2n}}{\partial x} & \cdots & \dfrac{\partial y_{mn}}{\partial x} \end{bmatrix} \in \mathbb{R}^{n\times m}, \tag{21}$$

where $\mathbf{Y} \in \mathbb{R}^{m\times n}$.

2.2.3 Scalar-matrix derivative (gradient matrix)

The partial derivative of $\dfrac{\partial y}{\partial \mathbf{X}}$ (usually called gradient matrix) is given by

$$\left[\frac{\partial y}{\partial \mathbf{X}}\right]_{\text{Den}} = \begin{bmatrix} \dfrac{\partial y}{\partial x_{11}} & \dfrac{\partial y}{\partial x_{12}} & \cdots & \dfrac{\partial y}{\partial x_{1n}} \\ \dfrac{\partial y}{\partial x_{21}} & \dfrac{\partial y}{\partial x_{22}} & \cdots & \dfrac{\partial y}{\partial x_{2n}} \\ \vdots & \vdots & \ddots & \vdots \\ \dfrac{\partial y}{\partial x_{m1}} & \dfrac{\partial y}{\partial x_{m2}} & \cdots & \dfrac{\partial y}{\partial x_{mn}} \end{bmatrix} \in \mathbb{R}^{m\times n}, \tag{22}$$

where $\mathbf{X} \in \mathbb{R}^{m\times n}$.

2.2.4 Row vector-scalar and scalar-row vector derivatives

From these definitions, we can infer two nonobvious equalities that are rather useful when handling matrix differentiations. If we consider a special case of the gradient matrix, (22), when $m = 1 \therefore \mathbf{X} = \mathbf{x}^\top \in \mathbb{R}^{1\times n}$, where $\mathbf{x} \in \mathbb{R}^n$, we have that

$$\left[\frac{\partial \alpha}{\partial \mathbf{x}^\top}\right]_{\text{Den}} = \begin{bmatrix} \dfrac{\partial \alpha}{\partial x_1} & \dfrac{\partial \alpha}{\partial x_2} & \cdots & \dfrac{\partial \alpha}{\partial x_n} \end{bmatrix} \in \mathbb{R}^{1\times n}, \tag{23}$$

however, by using we definition from (20), it is also true to state that

$$\left[\frac{\partial \alpha}{\partial \mathbf{x}}\right]_{\text{Den}} = \begin{bmatrix} \dfrac{\partial \alpha}{\partial x_1} \\ \dfrac{\partial \alpha}{\partial x_2} \\ \vdots \\ \dfrac{\partial \alpha}{\partial x_n} \end{bmatrix} \in \mathbb{R}^n. \tag{24}$$

Therefore,

$$\boxed{\left[\frac{\partial \alpha}{\partial \mathbf{x}^\top}\right]_{\text{Den}} = \left[\frac{\partial \alpha}{\partial \mathbf{x}}\right]_{\text{Den}}^\top} \tag{25}$$

Similarly, from (21), when $m = 1 \therefore \mathbf{Y} = \mathbf{y}^\top \in \mathbb{R}^{1\times n}$, where $\mathbf{y} \in \mathbb{R}^n$, we have that

$$\left[\frac{\partial \mathbf{y}^\top}{\partial \alpha}\right]_{\text{Den}} = \begin{bmatrix} \frac{\partial y_1}{\partial \alpha} \\ \frac{\partial y_2}{\partial \alpha} \\ \cdots \\ \frac{\partial y_n}{\partial \alpha} \end{bmatrix} \in \mathbb{R}^n. \tag{26}$$

However, from (20), we also have that

$$\left[\frac{\partial \mathbf{y}}{\partial \alpha}\right]_{\text{Den}} = \begin{bmatrix} \frac{\partial y_1}{\partial \alpha} & \frac{\partial y_2}{\partial \alpha} & \cdots & \frac{\partial y_n}{\partial \alpha} \end{bmatrix} \in \mathbb{R}^{1\times n}. \tag{27}$$

Therefore,

$$\boxed{\left[\frac{\partial \mathbf{y}^\top}{\partial \alpha}\right]_{\text{Den}} = \left[\frac{\partial \mathbf{y}}{\partial \alpha}\right]_{\text{Den}}^\top.} \tag{28}$$

2.2.5 The Jacobian matrix for the denominator layout

The Jacobian matrix in denominator notation is given by

$$\mathbf{J} = \left[\frac{\partial f(\mathbf{x})}{\partial \mathbf{x}}^\top\right]_{\text{Den}} \tag{29}$$

$$= \begin{bmatrix} \frac{\partial f_1(\mathbf{x})}{\partial \mathbf{x}}^\top \\ \frac{\partial f_2(\mathbf{x})}{\partial \mathbf{x}}^\top \\ \vdots \\ \frac{\partial f_m(\mathbf{x})}{\partial \mathbf{x}}^\top \end{bmatrix} \tag{30}$$

$$= \begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \frac{\partial f_2}{\partial x_1} & \frac{\partial f_2}{\partial x_2} & \cdots & \frac{\partial f_2}{\partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial f_m}{\partial x_1} & \frac{\partial f_m}{\partial x_2} & \cdots & \frac{\partial f_m}{\partial x_n} \end{bmatrix} \in \mathbb{R}^{m\times n}, \tag{31}$$

where $f : \mathbb{R}^n \to \mathbb{R}^m$ and $f = (f_1, f_2, \ldots, f_m)$, being $f_i : \mathbb{R} \to \mathbb{R}$ for $1 \leq i \leq m$.

2.2.6 The Hessian matrix for the denominator layout

The Hessian matrix in the denominator layout is given by

$$\mathbf{H} = \left[\frac{\partial^2 f(\mathbf{x})}{\partial \mathbf{x} \partial \mathbf{x}^\top}\right]_{\text{Den}} = \left[\frac{\partial}{\partial \mathbf{x}}\left(\frac{\partial f(\mathbf{x})}{\partial \mathbf{x}^\top}\right)\right]_{\text{Den}} = \begin{bmatrix} \frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_1 \partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_1 \partial x_n} \\ \frac{\partial^2 f}{\partial x_2 \partial x_1} & \frac{\partial^2 f}{\partial x_2^2} & \cdots & \frac{\partial^2 f}{\partial x_2 \partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 f}{\partial x_n \partial x_1} & \frac{\partial^2 f}{\partial x_n \partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_n^2} \end{bmatrix} \in \mathbb{R}^{n \times n}, \tag{32}$$

where $f : \mathbb{R}^n \to \mathbb{R}$ and $\mathbf{x} \in \mathbb{R}^n$. In the denominator layout, the visualization of the Hessian matrix is straightforward as it follows the denominator shape. It is important to point out that some authors [16, 17] prefer to denote the Hessian matrix as

$$\mathbf{H} = \left[\frac{\partial^2 f(\mathbf{x})}{\partial \mathbf{x}^2}\right]_{\text{Den}}, \tag{33}$$

where element $\partial \mathbf{x}^2$ is merely a shorthand for $\partial \mathbf{x} \partial \mathbf{x}^\top$. However, it might not be recommended as it makes the notation even more confusing.

2.3 Comparative between Jacobian and Hessian formulations

As you may have noticed,

$$\left[\frac{\partial \mathbf{y}}{\partial \mathbf{x}}\right]_{\text{Num}} = \left[\frac{\partial \mathbf{y}}{\partial \mathbf{x}}\right]_{\text{Den}}^\top, \tag{34}$$

$$\left[\frac{\partial \mathbf{Y}}{\partial x}\right]_{\text{Num}} = \left[\frac{\partial \mathbf{Y}}{\partial x}\right]_{\text{Den}}^\top, \tag{35}$$

$$\left[\frac{\partial y}{\partial \mathbf{X}}\right]_{\text{Num}} = \left[\frac{\partial y}{\partial \mathbf{X}}\right]_{\text{Den}}^\top, \tag{36}$$

$$\left[\frac{\partial \alpha}{\partial \mathbf{x} \partial \mathbf{x}^\top}\right]_{\text{Num}} = \left[\frac{\partial \alpha}{\partial \mathbf{x}^2}\right]_{\text{Den}}, \tag{37}$$

where that last equation can be inferred from the discussion about the Hessian matrix.

That is the difference when you try to differentiate without paying attention to which representation the author adopted. The good news is that, as long as you differentiate it correctly, you can switch between the Jacobian and Hessian formulations by simply transposing the final result. Fortunately, the denominator layout is the most adopted by authors from areas related to Signal Processing, Machine Learning, and Optimization Theory. That is why we will focus on the denominator layout hereafter (the notation $[\cdot]_{\text{Den}}$ will be dropped out since we do not need it anymore).

By adopting the denominator layout, keep in mind that:

- $\frac{\partial \mathbf{y}}{\partial \mathbf{x}}$ will yield a matrix.
- $\frac{\partial \mathbf{Y}}{\partial x}$ will yield a matrix.
- $\frac{\partial x}{\partial \mathbf{X}}$ will yield a matrix.
- $\frac{\partial y}{\partial \mathbf{x}}$ will yield a vector.
- $\frac{\partial \mathbf{y}}{\partial x}$ will yield a $1 \times n$ matrix ("row vector").

- $\mathbf{J} = \dfrac{\partial f(\mathbf{x})}{\partial \mathbf{x}}^\top$.
- $\mathbf{H} = \dfrac{\partial^2 f(\mathbf{x})}{\partial \mathbf{x} \partial \mathbf{x}^\top}$.
- $\dfrac{\partial \alpha}{\partial \mathbf{x}^\top} = \dfrac{\partial \alpha}{\partial \mathbf{x}}^\top$ for $\alpha \in \mathbb{R}$.
- $\dfrac{\partial \mathbf{y}^\top}{\partial \alpha} = \dfrac{\partial \mathbf{y}}{\partial \alpha}^\top$ for $\alpha \in \mathbb{R}$.

2.4 Notations not widely agreed upon

Expressions such as $\dfrac{\partial \mathbf{y}}{\partial \mathbf{X}}$, $\dfrac{\partial \mathbf{Y}}{\partial \mathbf{X}}$, or $\dfrac{\partial \mathbf{Y}}{\partial \mathbf{x}}$ have no agreement for both numerator and denominator layouts. It is possible, however, to define the matrix-matrix derivative for both representations. The problem is that some authors define it in the most intuitive manner: For $\dfrac{\partial \mathbf{Y}}{\partial \mathbf{X}}$, element in the position (i, j) is $\partial y_{ij} / \partial x_{ij}$. However, as we saw, it is inconsistent for both formulations as the Jacobian (Hessian) formulation must lay out its denominator (numerator) in its transposed shape. Therefore, for a consistent numerator layout, we would have

$$\left[\frac{\partial \mathbf{Y}}{\partial \mathbf{X}}\right]_{\text{Num}} = \begin{bmatrix} \dfrac{\partial y_{1,1}}{\partial x_{1,1}} & \dfrac{\partial y_{1,2}}{\partial x_{2,1}} & \cdots & \dfrac{\partial y_{1,m}}{\partial x_{m,1}} \\ \dfrac{\partial y_{2,1}}{\partial x_{1,2}} & \dfrac{\partial y_{2,2}}{\partial x_{2,2}} & \cdots & \dfrac{\partial y_{2,n}}{\partial x_{m,2}} \\ \vdots & \vdots & \ddots & \vdots \\ \dfrac{\partial y_{n,1}}{\partial x_{1,n}} & \dfrac{\partial y_{n,2}}{\partial x_{2,n}} & \cdots & \dfrac{\partial y_{n,m}}{\partial x_{m,n}} \end{bmatrix} \in \mathbb{R}^{n \times m}, \tag{38}$$

and for a consistent denominator layout, we would have

$$\left[\frac{\partial \mathbf{Y}}{\partial \mathbf{X}}\right]_{\text{Den}} = \begin{bmatrix} \dfrac{\partial y_{1,1}}{\partial x_{1,1}} & \dfrac{\partial y_{2,1}}{\partial x_{1,2}} & \cdots & \dfrac{\partial y_{n,1}}{\partial x_{1,n}} \\ \dfrac{\partial y_{1,2}}{\partial x_{2,1}} & \dfrac{\partial y_{2,2}}{\partial x_{2,2}} & \cdots & \dfrac{\partial y_{n,2}}{\partial x_{2,n}} \\ \vdots & \vdots & \ddots & \vdots \\ \dfrac{\partial y_{1,m}}{\partial x_{m,1}} & \dfrac{\partial y_{2,m}}{\partial x_{m,2}} & \cdots & \dfrac{\partial y_{n,m}}{\partial x_{m,n}} \end{bmatrix} \in \mathbb{R}^{m \times n}, \tag{39}$$

where $\mathbf{X} \in \mathbb{R}^{m \times n}$ and $\mathbf{Y} \in \mathbb{R}^{n \times m}$. For expressions $\dfrac{\partial \mathbf{y}}{\partial \mathbf{X}}$ and $\dfrac{\partial \mathbf{Y}}{\partial \mathbf{x}}$ can the derived by taking them as special cases of the matrix-matrix formulation. Notwithstanding, keep in mind that both equations are not standard and their usage must be acknowledged by the reader.

2.5 On the ambiguity over the notation ∇^2

It is common for some authors, such as Simon Haykin [10] and Bishop [9], to denote the gradient vector and the Hessian matrix as

$$\mathbf{g} = \nabla f = \frac{\partial f}{\partial \mathbf{x}} \tag{40}$$

and

$$\mathbf{H} = \nabla^2 f = \frac{\partial^2 f(\mathbf{x})}{\partial \mathbf{x} \partial \mathbf{x}^\top}, \tag{41}$$

respectively, where

$$\nabla = \begin{bmatrix} \frac{\partial}{\partial x_1} \\ \frac{\partial}{\partial x_2} \\ \vdots \\ \frac{\partial}{\partial x_n} \end{bmatrix} \tag{42}$$

is the gradient vector. Therefore, the symbol $\nabla^2 f$ means that the gradient vector is applied twice to the scalar function f: the first is a scalar-vector differentiation; the second is a vector-vector differentiation. It is important to point out, however, that the notation ∇^2 might be ambiguous since it is also used in Vector Calculus to denote another operator: the Laplacian.

Vector Calculus is another branch of Mathematics that deals with differentiations and integrals of vector fields in a three-dimensional space. It is used in different areas of Physics and Engineering, such as electromagnetic fields, fluid dynamics, etc. In Vector Calculus,

$$\nabla^2 f = \nabla \cdot \nabla f = \frac{\partial f}{\partial x_1} + \frac{\partial f}{\partial x_2} + \dots \frac{\partial f}{\partial x_n} \tag{43}$$

denotes the so-called Laplacian operator, where $\cdot$ is the dot or inner product. The equations (41) and (43) have the same notations but refer to different operations. That is a problem! Since Vector Calculus is far enough from Matrix Calculus in many applications, the notation $\nabla^2 f$ is used in both fields without problems. However, when this is not the case, the notation $\nabla^2 f$ must be avoided to denote the Hessian matrix. Other authors, such as Bishop [9], use $\nabla\nabla f$, which might be a way out to disambiguate it.

3 Identities

We need to be cautious when applying the matrix differentiation identities since the element orders matter. For instance, for scalar elements, the product rule may be written as either $(fg)' = f'g + g'f$ or $(fg)' = gf' + fg'$. In Matrix Calculus, we do not have such a privilege.

3.1 Chain rule

3.1.1 Univariate functions

For scalar elements, the chain rule is given by

$$\frac{\partial w}{\partial v} = \frac{\partial w}{\partial x}\frac{\partial x}{\partial y}\frac{\partial y}{\partial v}. \tag{44}$$

Similarly, in matrix notation, we have

$$\frac{\partial \mathbf{w}}{\partial \mathbf{v}} = \frac{\partial \mathbf{y}}{\partial \mathbf{v}}\frac{\partial \mathbf{x}}{\partial \mathbf{y}}\frac{\partial \mathbf{w}}{\partial \mathbf{x}}, \tag{45}$$

where $\mathbf{x} \in \mathbb{R}^n, \mathbf{y} \in \mathbb{R}^m, \mathbf{v} \in \mathbb{R}^p$, and $\mathbf{w} \in \mathbb{R}^q$. In this expression, $\mathbf{w}$ depends on $\mathbf{x}$, $\mathbf{x}$ depends on $\mathbf{y}$, and $\mathbf{y}$ depends on $\mathbf{z}$. The number of elements in the chain rule can be increased indiscriminately. The main point here is that *the chain rule in Matrix Calculus notation must be placed backwards when compared with the standard chain rule of scalar elements.*

3.1.2 Multivariate functions

In the previous section, we had a case where $\mathbf{w}$ depends on $\mathbf{x}$, which depends on $\mathbf{y}$, which depends on $\mathbf{v}$. If $\mathbf{w} = f(\mathbf{x}), \mathbf{x} = g(\mathbf{y})$, and $\mathbf{y} = h(\mathbf{v})$, then f, g, and h are functions of one variable, also called univariate functions. However, we might find a situation where $\mathbf{w} = f(\mathbf{x}, \mathbf{y})$ is a function of two (or more) variables.

For scalar elements, we can find partial derivatives of multivariate functions by considering that $w = f(x, y)$ is differentiable on x and y. The chain rule becomes

$$\frac{\partial w}{\partial z} = \frac{\partial w}{\partial x}\frac{\partial x}{\partial z} + \frac{\partial w}{\partial y}\frac{\partial y}{\partial z}. \tag{46}$$

Similarly, for Matrix Calculus notation, we have

$$\frac{\partial \mathbf{w}}{\partial \mathbf{z}} = \frac{\partial \mathbf{x}}{\partial \mathbf{z}}\frac{\partial \mathbf{w}}{\partial \mathbf{x}} + \frac{\partial \mathbf{y}}{\partial \mathbf{z}}\frac{\partial \mathbf{w}}{\partial \mathbf{y}} \tag{47}$$

Note the backward placement of each summation term. This expression can be used for an unrestricted number of variables.

3.2 Sum (or minus) rule

3.2.1 Vector-vector derivative

Let $\mathbf{x}, \mathbf{y} \in \mathbb{R}^m$ and $a, b \in \mathbb{R}$, where $\mathbf{x}$ and $\mathbf{y}$ depend on $\mathbf{w} \in \mathbb{R}^n$, but a and b do not. Therefore,

$$\frac{\partial (a\mathbf{x} \pm b\mathbf{y})}{\partial \mathbf{w}} = a\frac{\partial \mathbf{x}}{\partial \mathbf{w}} \pm b\frac{\partial \mathbf{y}}{\partial \mathbf{w}} \tag{48}$$

3.2.2 Matrix-scalar derivative

Another is when you have

$$\frac{\partial \left(a\mathbf{X} \pm b\mathbf{Y}\right)}{\partial \alpha}, \tag{49}$$

where $\mathbf{X}, \mathbf{Y} \in \mathbb{R}^{m\times n}$ depend on $\alpha \in \mathbb{R}$. The solution is

$$\frac{\partial (a\mathbf{X} \pm b\mathbf{Y})}{\partial \alpha} = a\frac{\partial \mathbf{X}}{\partial \alpha} \pm b\frac{\partial \mathbf{Y}}{\partial \alpha}. \tag{50}$$

3.2.3 Scalar-matrix derivative

The scalar-matrix derivative has a similar result, i.e.,

$$\frac{\partial \left(ax \pm by\right)}{\partial \mathbf{W}} = a\frac{\partial x}{\partial \mathbf{W}} \pm b\frac{\partial y}{\partial \mathbf{W}}, \tag{51}$$

where $x, y \in \mathbb{R}$ depend on $\mathbf{W} \in \mathbb{R}^{m\times n}$, but $a, b \in \mathbb{R}$ do not.

3.3 Product rule

3.3.1 Vector-vector derivative

Let $w \in \mathbb{R}$ and $\mathbf{v} \in \mathbb{R}^m$, where both depend on $\mathbf{x} \in \mathbb{R}^n$. Then,

$$\frac{\partial w\mathbf{v}}{\partial \mathbf{x}} = w\frac{\partial \mathbf{v}}{\partial \mathbf{x}} + \frac{\partial w}{\partial \mathbf{x}}\mathbf{v}^\top. \tag{52}$$

Note that is not possible to apply the product rule when you have $\mathbf{W}\mathbf{v}$, where $\mathbf{W} \in \mathbb{R}^{n\times m}$ also depends on $\mathbf{x}$. If you tried, you would get $\partial \mathbf{W}/\partial \mathbf{x}$, which there is no consensus about this layout (vide Section 2.4).

3.3.2 Scalar-vector derivative

Another possibility of applying the product rule is when you have $\mathbf{w}^\top\mathbf{v}$, where $\mathbf{w} \in \mathbb{R}^m$ also depends on $\mathbf{x} \in \mathbb{R}^n$. In this case, the dot product is given by

$$\frac{\partial \mathbf{w}^\top\mathbf{v}}{\partial \mathbf{x}} = \frac{\partial \mathbf{v}}{\partial \mathbf{x}}\mathbf{w} + \frac{\partial \mathbf{w}}{\partial \mathbf{x}}\mathbf{v}. \tag{53}$$

3.3.3 Scalar-matrix derivative

It is still possible to apply the product rule to

$$\frac{\partial wv}{\partial \mathbf{X}}, \tag{54}$$

where $w, v \in \mathbb{R}$ depend on $\mathbf{X} \in \mathbb{R}^{m \times n}$. In this case, we have

$$\frac{\partial wv}{\partial \mathbf{X}} = w\frac{\partial v}{\partial \mathbf{X}} + v\frac{\partial w}{\partial \mathbf{X}}. \tag{55}$$

3.3.4 Matrix-scalar derivative

The last case is when you have

$$\frac{\partial \mathbf{WV}}{\partial \alpha}, \tag{56}$$

where both $\mathbf{W} \in \mathbb{R}^{m \times p}$ and $\mathbf{V} \in \mathbb{R}^{p \times n}$ depend on $\alpha \in \mathbb{R}$. In this case, we have

$$\frac{\partial \mathbf{WV}}{\partial \alpha} = \frac{\partial \mathbf{V}}{\partial \alpha}\mathbf{W}^\top + \mathbf{V}^T\frac{\mathbf{W}}{\partial \alpha} \tag{57}$$

4 Unconstrained optimization algorithms

In this section, we use Matrix Calculus to derive unconstrained optimization algorithms. The key idea is to show how Matrix Calculus allows us to conceive simple and elegant expressions, without resorting to nonmatrix notation. However, we are not interested in the theoretical aspects of such algorithms, e.g., necessary and sufficient conditions for unconstrained minimization, Lipschitz continuity, Kantorovich theorem, etc. For these topics, the reader is encouraged to consult [18].

4.1 Taylor Series

(TODO: see Tasinaffo material)

4.2 The Steepest Descent algorithm

4.3 Newton's Method

4.4 The Gauss-Newton Method

4.5 The Levenberg-Marquardt Method.

(See [19], in the unconstrained optimization methods; [20]; [21])

Part II
Manual Solution

5 Differentiation solutions

We have two ways to solve matrix differentiation:

1. Performing element-by-element operations in matrices and vectors;
2. Preserving the Matrix Calculus notation, performing operations on the whole matrix/vector and, eventually, using some identities.

The latter is usually more straightforward and less toilsome than the former and is therefore preferable. In this section, we will solve the expressions in both ways.

5.1 $\dfrac{\partial \mathbf{A}\mathbf{x}}{\partial \mathbf{x}} = \mathbf{A}^\top$

Let $\mathbf{A} \in \mathbb{R}^{m\times n}$ and $\mathbf{x} \in \mathbb{R}^n$, in which $\mathbf{A}$ does not depend on $\mathbf{x}$, we have that:

$$\frac{\partial \mathbf{A}\mathbf{x}}{\partial \mathbf{x}} = \frac{\partial}{\partial \mathbf{x}}\left(\begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}\right) \tag{58}$$

$$= \frac{\partial}{\partial \mathbf{x}}\left(\begin{bmatrix} \sum_{j=1}^n a_{1j}x_j \\ \sum_{j=1}^n a_{2j}x_j \\ \vdots \\ \sum_{j=1}^n a_{mj}x_j \end{bmatrix}^\top\right) \tag{59}$$

$$= \begin{bmatrix} \frac{\partial}{\partial \mathbf{x}}\left(\sum_{j=1}^n a_{1j}x_j\right) & \frac{\partial}{\partial \mathbf{x}}\left(\sum_{j=1}^n a_{2j}x_j\right) & \dots & \frac{\partial}{\partial \mathbf{x}}\left(\sum_{j=1}^n a_{mj}x_j\right) \end{bmatrix} \tag{60}$$

Since a scalar-vector derivative is represented by a vector, we have that

$$\frac{\partial \mathbf{A}\mathbf{x}}{\partial \mathbf{x}} = \begin{bmatrix} \frac{\partial}{\partial x_1}\left(\sum_{j=1}^n a_{1j}x_j\right) & \frac{\partial}{\partial x_1}\left(\sum_{j=1}^n a_{21j}x_j\right) & \dots & \frac{\partial}{\partial x_1}\left(\sum_{j=1}^n a_{mj}x_j\right) \\ \frac{\partial}{\partial x_2}\left(\sum_{j=1}^n a_{1j}x_j\right) & \frac{\partial}{\partial x_2}\left(\sum_{j=1}^n a_{21j}x_j\right) & \dots & \frac{\partial}{\partial x_2}\left(\sum_{j=1}^n a_{mj}x_j\right) \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial}{\partial x_n}\left(\sum_{j=1}^n a_{1j}x_j\right) & \frac{\partial}{\partial x_n}\left(\sum_{j=1}^n a_{21j}x_j\right) & \dots & \frac{\partial}{\partial x_n}\left(\sum_{j=1}^n a_{mj}x_j\right) \end{bmatrix} \tag{61}$$

$$= \begin{bmatrix} a_{11} & a_{21} & \dots & a_{n1} \\ a_{12} & a_{22} & \dots & a_{n2} \\ \vdots & \vdots & \ddots & \vdots \\ a_{1m} & a_{2m} & \dots & a_{nm} \end{bmatrix} \tag{62}$$

$$\boxed{\frac{\partial \mathbf{A}\mathbf{x}}{\partial \mathbf{x}} = \mathbf{A}^\top \in \mathbb{R}^{n\times m}} \tag{63}$$

5.2 $\dfrac{\partial \mathbf{x}^\top \mathbf{A}}{\partial \mathbf{x}} = \mathbf{A}$

(TODO)

5.3 $\dfrac{\partial \mathbf{A}\mathbf{x}}{\partial \mathbf{v}} = \dfrac{\partial \mathbf{x}}{\partial \mathbf{v}}\mathbf{A}^\top$

Let $\mathbf{x} \in \mathbb{R}^n$, $\mathbf{v} \in \mathbb{R}^p$ and $\mathbf{A} \in \mathbb{R}^{m\times n}$, where $\mathbf{x}$ depends on $\mathbf{v}$, but $\mathbf{A}$ does not. Then

$$\frac{\partial \mathbf{A}\mathbf{x}}{\partial \mathbf{v}} = \begin{bmatrix} \frac{\partial}{\partial \mathbf{v}}\sum_{i=1}^n a_{1i}x_i & \frac{\partial}{\partial \mathbf{v}}\sum_{i=1}^n a_{2i}x_i & \cdots & \frac{\partial}{\partial \mathbf{v}}\sum_{i=1}^n a_{mi}x_i \end{bmatrix} \tag{64}$$

$$= \begin{bmatrix} \sum_{i=1}^n a_{1i}\frac{\partial x_i}{\partial \mathbf{v}} & \sum_{i=1}^n a_{2i}\frac{\partial x_i}{\partial \mathbf{v}} & \cdots & \sum_{i=1}^n a_{mi}\frac{\partial x_i}{\partial \mathbf{v}} \end{bmatrix} \tag{65}$$

$$= \underbrace{\begin{bmatrix} \frac{\partial x_1}{\partial \mathbf{v}} & \frac{\partial x_2}{\partial \mathbf{v}} & \cdots & \frac{\partial x_n}{\partial \mathbf{v}} \end{bmatrix}}_{p\times n}\underbrace{\begin{bmatrix} a_{11} & a_{21} & \dots & a_{m1} \\ a_{12} & a_{22} & \dots & a_{m2} \\ \vdots & \vdots & \ddots & \vdots \\ a_{1n} & a_{2n} & \dots & a_{mn} \end{bmatrix}}_{n\times m} \tag{66}$$

$$\boxed{\frac{\partial \mathbf{A}\mathbf{x}}{\partial \mathbf{v}} = \frac{\partial \mathbf{x}}{\partial \mathbf{v}}\mathbf{A}^\top \in \mathbb{R}^{p \times m}} \tag{67}$$

Observe that this result is equivalent to applying the chain rule (cf. (45)), that is,

$$\frac{\partial \mathbf{A}\mathbf{x}}{\partial \mathbf{v}} = \frac{\partial \mathbf{x}}{\partial \mathbf{v}}\frac{\partial \mathbf{A}\mathbf{x}}{\partial \mathbf{x}} = \frac{\partial \mathbf{x}}{\partial \mathbf{v}}\mathbf{A}^\top. \tag{68}$$

5.4 $\frac{\partial \mathbf{a}^\top \mathbf{x}}{\partial \mathbf{x}} = \mathbf{a}$

Let $\mathbf{a}, \mathbf{x} \in \mathbb{R}^n$, in which $\mathbf{a}$ does not depend on $\mathbf{x}$. You can derive the derivative for this inner product by considering that $\mathbf{a}^\top$ is actually a $1 \times n$ matrix that transforms a $\mathbb{R}^n$ space into $\mathbb{R}$, and we already know that $\partial \mathbf{A}\mathbf{x}/\partial \mathbf{x} = \mathbf{A}^\top$. Thus,

$$\frac{\partial \mathbf{a}^\top \mathbf{x}}{\partial \mathbf{x}} = \mathbf{a}^{\top\top} = \mathbf{a}. \tag{69}$$

Even though, if you want the step-by-step, here it is:

$$\begin{aligned}
\frac{\partial \mathbf{a}^\top \mathbf{x}}{\partial \mathbf{x}} &= \frac{\partial}{\partial \mathbf{x}}\left(\begin{bmatrix} a_1 & a_2 & \dots & a_n \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}\right) &(70)\\
&= \frac{\partial}{\partial \mathbf{x}}\left(\sum_{i=1}^{n} a_i x_i\right) &(71)\\
&= \begin{bmatrix} \frac{\partial}{\partial x_1}\left(\sum_{i=1}^{n} a_i x_i\right) \\ \frac{\partial}{\partial x_2}\left(\sum_{i=1}^{n} a_i x_i\right) \\ \vdots \\ \frac{\partial}{\partial x_n}\left(\sum_{i=1}^{n} a_i x_i\right) \end{bmatrix} = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix} &(72)
\end{aligned}$$

$$\boxed{\frac{\partial \mathbf{a}^\top \mathbf{x}}{\partial \mathbf{x}} = \mathbf{a} \in \mathbb{R}^n} \tag{73}$$

5.5 $\frac{\partial \mathbf{x}^\top \mathbf{a}}{\partial \mathbf{x}} = \mathbf{a}$

This one can be solved quickly by noticing that $\mathbf{x}^\top \mathbf{a} = \mathbf{a}^\top \mathbf{x}$. Hence,

$$\frac{\partial \mathbf{x}^\top \mathbf{a}}{\partial \mathbf{x}} = \frac{\partial \mathbf{a}^\top \mathbf{x}}{\partial \mathbf{x}} = \mathbf{a} \tag{74}$$

Nevertheless, here is the step-by-step:

$$\frac{\partial \mathbf{x}^\top \mathbf{a}}{\partial \mathbf{x}} = \frac{\partial}{\partial \mathbf{x}} \left(\begin{bmatrix} x_1 & x_2 & \dots & x_n \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix} \right) \tag{75}$$

$$= \frac{\partial}{\partial \mathbf{x}} \left(\sum_{i=1}^{n} x_i a_i \right) \tag{76}$$

$$= \begin{bmatrix} \frac{\partial}{\partial x_1} \left(\sum_{i=1}^{n} x_i a_i \right) \\ \frac{\partial}{\partial x_2} \left(\sum_{i=1}^{n} x_i a_i \right) \\ \vdots \\ \frac{\partial}{\partial x_n} \left(\sum_{i=1}^{n} x_i a_i \right) \end{bmatrix} = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix} \tag{77}$$

$$\boxed{\frac{\partial \mathbf{x}^\top \mathbf{a}}{\partial \mathbf{x}} = \mathbf{a} \in \mathbb{R}^n} \tag{78}$$

5.6 $\frac{\partial \mathbf{a}^\mathsf{H} \mathbf{x}}{\partial \mathbf{x}} = \mathbf{a}^*$

Let $\mathbf{a} \in \mathbb{C}^n$ and $\mathbf{x} \in \mathbb{R}^n$, in which $\mathbf{a}$ does not depend on $\mathbf{x}$. Once again, we could say that

$$\frac{\partial \mathbf{a}^\mathsf{H} \mathbf{x}}{\partial \mathbf{x}} = \mathbf{a}^{\mathsf{H}^\top} = \mathbf{a}^*, \tag{79}$$

where $\cdot^*$ denotes the conjugate operator.

Nevertheless, here is the step-by-step:

$$\frac{\partial \mathbf{a}^\mathsf{H} \mathbf{x}}{\partial \mathbf{x}} = \frac{\partial}{\partial \mathbf{x}} \left(\begin{bmatrix} a_1^* & a_2^* & \dots & a_n^* \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \right) = \frac{\partial}{\partial \mathbf{x}} \left(\sum_{i=1}^{n} a_i^* x_i \right) \tag{80}$$

$$\tag{81}$$

Since a scalar-vector derivative is represented by a vector, we have that

$$\frac{\partial \mathbf{a}^\mathsf{H} \mathbf{x}}{\partial \mathbf{x}} = \begin{bmatrix} \frac{\partial}{\partial x_1} \left(\sum_{i=1}^{n} a_i^* x_i \right) \\ \frac{\partial}{\partial x_2} \left(\sum_{i=1}^{n} a_i^* x_i \right) \\ \vdots \\ \frac{\partial}{\partial x_n} \left(\sum_{i=1}^{n} a_i^* x_i \right) \end{bmatrix} = \begin{bmatrix} a_1^* \\ a_2^* \\ \vdots \\ a_n^* \end{bmatrix} \tag{82}$$

$$\boxed{\frac{\partial \mathbf{a}^\mathsf{H} \mathbf{x}}{\partial \mathbf{x}} = \mathbf{a}^* \in \mathbb{C}^n} \tag{83}$$

5.7 $\dfrac{\partial \mathbf{z}^{\mathsf{H}}\mathbf{a}}{\partial \mathbf{z}} = \mathbf{0}$

Let us define $\mathbf{z}, \mathbf{a} \in \mathbb{C}^n$, where $\mathbf{a}$ does not depend on $\mathbf{z}$. Since $\mathbf{z}^{\mathsf{H}}\mathbf{a} \neq \mathbf{a}^{\mathsf{H}}\mathbf{z}$, we have no choice but derive it:

$$\frac{\partial \mathbf{z}^{\mathsf{H}}\mathbf{a}}{\partial \mathbf{z}} = \frac{\partial}{\partial \mathbf{z}}\left(\begin{bmatrix} z_1^* & z_2^* & \dots & z_n^* \end{bmatrix}\begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix}\right) \tag{84}$$

$$= \frac{\partial}{\partial \mathbf{z}}\left(\sum_{i=1}^{n} z_i^* a_i\right) \tag{85}$$

$$= \begin{bmatrix} \frac{\partial}{\partial z_1}\left(\sum_{i=1}^{n} z_i^* a_i\right) \\ \frac{\partial}{\partial z_2}\left(\sum_{i=1}^{n} z_i^* a_i\right) \\ \vdots \\ \frac{\partial}{\partial z_n}\left(\sum_{i=1}^{n} z_i^* a_i\right) \end{bmatrix} \tag{86}$$

$$= \begin{bmatrix} \frac{\partial z_1^*}{\partial z_1} \\ \frac{\partial z_2^*}{\partial z_2} \\ \vdots \\ \frac{\partial z_n^*}{\partial z_n} \end{bmatrix}. \tag{87}$$

By recalling that $\dfrac{\partial z^*}{\partial z} = 0$ [6], we have that

$$\frac{\partial \mathbf{z}^{\mathsf{H}}\mathbf{a}}{\partial \mathbf{z}} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \tag{88}$$

$$\boxed{\frac{\partial \mathbf{z}^{\mathsf{H}}\mathbf{a}}{\partial \mathbf{z}} = \mathbf{0} \in \mathbb{C}^n} \tag{89}$$

where $\mathbf{0}$ is the zero vector.

5.8 $\dfrac{\partial \mathbf{z}^{\mathsf{H}}\mathbf{a}}{\partial \mathbf{z}^*} = \mathbf{a}$

(TODO) [6]

5.9 $\dfrac{\partial \mathbf{x}^\top \mathbf{y}}{\partial \mathbf{v}} = \dfrac{\partial \mathbf{x}}{\partial \mathbf{v}}\mathbf{y} + \dfrac{\partial \mathbf{y}}{\partial \mathbf{v}}\mathbf{x}$

Let $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ and $\mathbf{v} \in \mathbb{R}^m$. Where $\mathbf{x}$ and $\mathbf{y}$ depend on $\mathbf{v}$. Thus,

$$\frac{\partial \mathbf{x}^\top \mathbf{y}}{\partial \mathbf{v}} = \frac{\partial}{\partial \mathbf{v}} \sum_{i=1}^{n} y_i x_i \tag{90}$$

$$= \sum_{i=1}^{n} \frac{\partial y_i x_i}{\partial \mathbf{v}}. \tag{91}$$

Recalling that $(fg)' = f'g + g'f$, we have

$$\frac{\partial \mathbf{x}^\top \mathbf{y}}{\partial \mathbf{v}} = \sum_{i=1}^{n} x_i \frac{\partial y_i}{\partial \mathbf{v}} + \sum_{i=1}^{n} y_i \frac{\partial x_i}{\partial \mathbf{v}} \tag{92}$$

$$= \begin{bmatrix} \sum_{i=1}^{n} x_i \dfrac{\partial y_i}{\partial v_1} \\ \sum_{i=1}^{n} x_i \dfrac{\partial y_i}{\partial v_2} \\ \vdots \\ \sum_{i=1}^{n} x_i \dfrac{\partial y_i}{\partial v_m} \end{bmatrix} + \begin{bmatrix} \sum_{i=1}^{n} y_i \dfrac{\partial x_i}{\partial v_1} \\ \sum_{i=1}^{n} y_i \dfrac{\partial x_i}{\partial v_2} \\ \vdots \\ \sum_{i=1}^{n} y_i \dfrac{\partial x_i}{\partial v_m} \end{bmatrix} \tag{93}$$

$$= \begin{bmatrix} \dfrac{\partial y_1}{\partial v_1} & \dfrac{\partial y_2}{\partial v_1} & \cdots & \dfrac{\partial y_n}{\partial v_1} \\ \dfrac{\partial y_1}{\partial v_2} & \dfrac{\partial y_2}{\partial v_2} & \cdots & \dfrac{\partial y_n}{\partial v_2} \\ \vdots & \ddots & \vdots & \vdots \\ \dfrac{\partial y_1}{\partial v_m} & \dfrac{\partial y_2}{\partial v_m} & \cdots & \dfrac{\partial y_n}{\partial v_m} \end{bmatrix} \mathbf{x} \tag{94}$$

$$+ \begin{bmatrix} \dfrac{\partial x_1}{\partial v_1} & \dfrac{\partial x_2}{\partial v_1} & \cdots & \dfrac{\partial x_n}{\partial v_1} \\ \dfrac{\partial x_1}{\partial v_2} & \dfrac{\partial x_2}{\partial v_2} & \cdots & \dfrac{\partial x_n}{\partial v_2} \\ \vdots & \ddots & \vdots & \vdots \\ \dfrac{\partial x_1}{\partial v_m} & \dfrac{\partial x_2}{\partial v_m} & \cdots & \dfrac{\partial x_n}{\partial v_m} \end{bmatrix} \mathbf{y} \tag{95}$$

$$\boxed{\frac{\partial \mathbf{x}^\top \mathbf{y}}{\partial \mathbf{v}} = \frac{\partial \mathbf{x}}{\partial \mathbf{v}}\mathbf{y} + \frac{\partial \mathbf{y}}{\partial \mathbf{v}}\mathbf{x} \in \mathbb{R}^m} \tag{96}$$

Note that, if either $\mathbf{x}$ or $\mathbf{y}$ does not depend on $\mathbf{v}$, just disregard $\dfrac{\partial \mathbf{x}}{\partial \mathbf{v}}\mathbf{y}$ or $\dfrac{\partial \mathbf{y}}{\partial \mathbf{v}}\mathbf{x}$, respectively. When none depends on $\mathbf{v}$, the obvious result is the zero vector, $\mathbf{0}$. A simpler way to solve it is to apply the scalar-vector product rule (see (53)), that is,

$$\frac{\partial \mathbf{x}^\top \mathbf{y}}{\partial \mathbf{v}} = \frac{\partial \mathbf{x}}{\partial \mathbf{v}} \frac{\partial \mathbf{x}^\top \mathbf{y}}{\partial \mathbf{x}} + \frac{\partial \mathbf{y}}{\partial \mathbf{v}} \frac{\partial \mathbf{x}^\top \mathbf{y}}{\partial \mathbf{y}} = \frac{\partial \mathbf{x}}{\partial \mathbf{v}}\mathbf{y} + \frac{\partial \mathbf{y}}{\partial \mathbf{v}}\mathbf{x} \tag{97}$$

5.10 $\dfrac{\partial \mathbf{x}^\top \mathbf{x}}{\partial \mathbf{x}} = 2\mathbf{x}$

Let $\mathbf{x} \in \mathbb{R}^n$. Thus,

$$\frac{\partial \mathbf{x}^\top \mathbf{x}}{\partial \mathbf{x}} = \frac{\partial}{\partial \mathbf{x}} \sum_{i=1}^{n} x_i^2 \tag{98}$$

$$= \sum_{i=1}^{n} \frac{\partial x_i^2}{\partial \mathbf{x}} \tag{99}$$

$$= \begin{bmatrix} \sum_{i=1}^{n} \frac{\partial x_i^2}{\partial x_1} \\ \sum_{i=1}^{n} \frac{\partial x_i^2}{\partial x_2} \\ \vdots \\ \sum_{i=1}^{n} \frac{\partial x_i^2}{\partial x_m} \end{bmatrix} \tag{100}$$

$$= \begin{bmatrix} 2x_1 \\ 2x_2 \\ \vdots \\ 2x_n \end{bmatrix} \tag{101}$$

$$\boxed{\frac{\partial \mathbf{x}^\top \mathbf{x}}{\partial \mathbf{x}} = 2\mathbf{x} \in \mathbb{R}^n} \tag{102}$$

Note that this perfectly matches with the derivate of a quadratic scalar value, i.e., $\frac{\mathrm{d}x^2}{\mathrm{d}x} = 2x$.

5.11 $\frac{\partial \mathbf{x}^\top \mathbf{x}}{\partial \mathbf{v}} = 2\frac{\partial \mathbf{x}}{\partial \mathbf{v}}\mathbf{x}$

Let $\mathbf{x} \in \mathbb{R}^n$ and $\mathbf{v} \in \mathbb{R}^m$, where $\mathbf{x}$ depends on $\mathbf{v}$. Thus,

$$\frac{\partial \mathbf{x}^\top \mathbf{x}}{\partial \mathbf{v}} = \frac{\partial}{\partial \mathbf{v}} \sum_{i=1}^n x_i^2 \tag{103}$$

$$= \sum_{i=1}^n \frac{\partial x_i^2}{\partial \mathbf{v}} \tag{104}$$

$$= \begin{bmatrix} \sum_{i=1}^n \frac{\partial x_i^2}{\partial v_1} \\ \sum_{i=1}^n \frac{\partial x_i^2}{\partial v_2} \\ \vdots \\ \sum_{i=1}^n \frac{\partial x_i^2}{\partial v_m} \end{bmatrix} \tag{105}$$

$$= \begin{bmatrix} \sum_{i=1}^n 2x_i \frac{\partial x_i}{\partial v_1} \\ \sum_{i=1}^n 2x_i \frac{\partial x_i}{\partial v_2} \\ \vdots \\ \sum_{i=1}^n 2x_i \frac{\partial x_i}{\partial v_m} \end{bmatrix} \tag{106}$$

$$= 2 \begin{bmatrix} \frac{\partial x_1}{\partial \mathbf{v}} & \frac{\partial x_2}{\partial \mathbf{v}} & \dots & \frac{\partial x_n}{\partial \mathbf{v}} \end{bmatrix} \mathbf{x} \tag{107}$$

$$\boxed{\frac{\partial \mathbf{x}^\top \mathbf{x}}{\partial \mathbf{v}} = 2\frac{\partial \mathbf{x}}{\partial \mathbf{v}}\mathbf{x} \in \mathbb{R}^m} \tag{108}$$

Note that this solution could also be solved by the chain rule (cf. (45)) as follows

$$\boxed{\frac{\partial \mathbf{x}^\top \mathbf{x}}{\partial \mathbf{v}} = \frac{\partial \mathbf{x}}{\partial \mathbf{v}} \frac{\partial \mathbf{x}^\top \mathbf{x}}{\partial \mathbf{x}} = 2\frac{\partial \mathbf{x}}{\partial \mathbf{v}}\mathbf{x} \in \mathbb{R}^m} \tag{109}$$

5.12 $\frac{\partial \mathbf{x}^\top \mathbf{A} \mathbf{x}}{\partial \mathbf{x}} = \left(\mathbf{A}^\top + \mathbf{A}\right)\mathbf{x}$

Let $\mathbf{A} \in \mathbb{R}^{n\times n}$ and $\mathbf{x} \in \mathbb{R}^n$, in which $\mathbf{A}$ does not depend on $\mathbf{x}$. For the quadratic form, it follows that

$$\frac{\partial \mathbf{x}^\top \mathbf{A} \mathbf{x}}{\partial \mathbf{x}} = \frac{\partial}{\partial \mathbf{x}} \left(\begin{bmatrix} x_1 & x_2 & \dots & x_n \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \right) \tag{110}$$

$$= \frac{\partial}{\partial \mathbf{x}} \left(\begin{bmatrix} \sum_{i=1}^n x_i a_{i1} & \sum_{i=1}^n x_i a_{i2} & \dots & \sum_{i=1}^n x_i a_{in} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \right) \tag{111}$$

$$= \frac{\partial}{\partial \mathbf{x}} \left(\sum_{i=1}^n \sum_{j=1}^n a_{ij} x_i x_j \right). \tag{112}$$

Note that the element inside the parentheses is a scalar and that a scalar-vector derivative results in a vector, that is,

$$\frac{\partial \mathbf{x}^\top \mathbf{A}\mathbf{x}}{\partial \mathbf{x}} = \begin{bmatrix} \frac{\partial}{\partial x_1}\left(\sum_{i=1}^{n}\sum_{j=1}^{n} x_i a_{ij} x_j\right) \\ \frac{\partial}{\partial x_2}\left(\sum_{i=1}^{n}\sum_{j=1}^{n} x_i a_{ij} x_j\right) \\ \vdots \\ \frac{\partial}{\partial x_n}\left(\sum_{i=1}^{n}\sum_{j=1}^{n} x_i a_{ij} x_j\right) \end{bmatrix} \tag{113}$$

$$= \begin{bmatrix} 2x_1 a_{11} + \sum_{\substack{j=1 \\ j\neq 1}}^{n} a_{1j} x_j + \sum_{\substack{i=1 \\ i\neq 1}}^{n} a_{i1} x_i \\ 2x_2 a_{22} + \sum_{\substack{j=1 \\ j\neq 2}}^{n} a_{2j} x_j + \sum_{\substack{i=1 \\ i\neq 2}}^{n} a_{i2} x_i \\ \vdots \\ 2x_n a_{nn} + \sum_{\substack{j=1 \\ j\neq n}}^{n} a_{nj} x_j + \sum_{\substack{i=1 \\ i\neq n}}^{n} a_{in} x_i \end{bmatrix} \tag{114}$$

$$= \begin{bmatrix} \sum_{j=1}^{n} a_{1j} x_j \\ \sum_{j=1}^{n} a_{2j} x_j \\ \vdots \\ \sum_{j=1}^{n} a_{nj} x_j \end{bmatrix} + \begin{bmatrix} \sum_{i=1}^{n} a_{i1} x_i \\ \sum_{i=1}^{n} a_{i2} x_i \\ \vdots \\ \sum_{i=1}^{n} a_{in} x_i \end{bmatrix} \tag{115}$$

$$= \mathbf{A}^\top \mathbf{x} + \mathbf{A}\mathbf{x} \tag{116}$$

$$\boxed{\frac{\partial \mathbf{x}^\top \mathbf{A}\mathbf{x}}{\partial \mathbf{x}} = \left(\mathbf{A}^\top + \mathbf{A}\right)\mathbf{x} \in \mathbb{R}^n} \tag{117}$$

For the special case where $\mathbf{A}$ is symmetric, we obtain

$$\boxed{\frac{\partial \mathbf{x}^\top \mathbf{A}\mathbf{x}}{\partial \mathbf{x}} = 2\mathbf{A}\mathbf{x} \in \mathbb{R}^n} \tag{118}$$

5.13 $\frac{\partial \mathbf{x}^\top \mathbf{A}\mathbf{x}}{\partial \mathbf{v}} = \frac{\partial \mathbf{x}}{\partial \mathbf{v}}\left(\mathbf{A} + \mathbf{A}^\top\right)\mathbf{x}$

Let $\mathbf{v} \in \mathbb{R}^m, \mathbf{x} \in \mathbb{R}^n$ and $\mathbf{A} \in \mathbb{R}^{n\times n}$, where $\mathbf{x}$ depends on $\mathbf{v}$, but $\mathbf{A}$ does not. For the quadratic form, it follows that

$$\frac{\partial \mathbf{x}^\top \mathbf{A}\mathbf{x}}{\partial \mathbf{v}} = \frac{\partial}{\partial \mathbf{v}}\left(\begin{bmatrix} x_1 & x_2 & \dots & x_n \end{bmatrix}\begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}\right) \tag{119}$$

$$= \frac{\partial}{\partial \mathbf{v}}\left(\begin{bmatrix} \sum_{i=1}^n x_i a_{i1} & \sum_{i=1}^n x_i a_{i2} & \dots & \sum_{i=1}^n x_i a_{in} \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}\right) \tag{120}$$

$$= \frac{\partial}{\partial \mathbf{v}}\left(\sum_{i=1}^n \sum_{j=1}^n a_{ij} x_i x_j\right) \tag{121}$$

$$= \begin{bmatrix} \frac{\partial}{\partial v_1}\left(\sum_{i=1}^n \sum_{j=1}^n x_i a_{ij} x_j\right) \\ \frac{\partial}{\partial v_2}\left(\sum_{i=1}^n \sum_{j=1}^n x_i a_{ij} x_j\right) \\ \vdots \\ \frac{\partial}{\partial v_n}\left(\sum_{i=1}^n \sum_{j=1}^n x_i a_{ij} x_j\right) \end{bmatrix} \tag{122}$$

$$= \begin{bmatrix} \sum_{i=1}^n \sum_{j=1}^n a_{ij} \frac{\partial x_i x_j}{\partial v_1} \\ \sum_{i=1}^n \sum_{j=1}^n a_{ij} \frac{\partial x_i x_j}{\partial v_2} \\ \vdots \\ \sum_{i=1}^n \sum_{j=1}^n a_{ij} \frac{\partial x_i x_j}{\partial v_n} \end{bmatrix} \tag{123}$$

Recalling that $(fg)' = f'g + g'f$, we have that

$$\frac{\partial \mathbf{x}^\top \mathbf{A}\mathbf{x}}{\partial \mathbf{v}} = \begin{bmatrix} \sum_{i=1}^{n}\sum_{j=1}^{n} x_j a_{ij} \frac{\partial x_i}{\partial v_1} + \sum_{i=1}^{n}\sum_{j=1}^{n} x_i a_{ij} \frac{\partial x_j}{\partial v_1} \\ \sum_{i=1}^{n}\sum_{j=1}^{n} x_j a_{ij} \frac{\partial x_i}{\partial v_2} + \sum_{i=1}^{n}\sum_{j=1}^{n} x_i a_{ij} \frac{\partial x_j}{\partial v_2} \\ \vdots \\ \sum_{i=1}^{n}\sum_{j=1}^{n} x_j a_{ij} \frac{\partial x_i}{\partial v_n} + \sum_{i=1}^{n}\sum_{j=1}^{n} x_i a_{ij} \frac{\partial x_j}{\partial v_n} \end{bmatrix} \tag{124}$$

$$= \begin{bmatrix} \sum_{i=1}^{n}\sum_{j=1}^{n} x_j a_{ij} \frac{\partial x_i}{\partial v_1} \\ \sum_{i=1}^{n}\sum_{j=1}^{n} x_j a_{ij} \frac{\partial x_i}{\partial v_2} \\ \vdots \\ \sum_{i=1}^{n}\sum_{j=1}^{n} x_j a_{ij} \frac{\partial x_i}{\partial v_n} \end{bmatrix} + \begin{bmatrix} \sum_{i=1}^{n}\sum_{j=1}^{n} x_i a_{ij} \frac{\partial x_j}{\partial v_1} \\ \sum_{i=1}^{n}\sum_{j=1}^{n} x_i a_{ij} \frac{\partial x_j}{\partial v_2} \\ \vdots \\ \sum_{i=1}^{n}\sum_{j=1}^{n} x_i a_{ij} \frac{\partial x_j}{\partial v_n} \end{bmatrix} \tag{125}$$

$$= \begin{bmatrix} \frac{\partial x_1}{\partial \mathbf{v}} & \frac{\partial x_2}{\partial \mathbf{v}} & \dots & \frac{\partial x_n}{\partial \mathbf{v}} \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} +$$

$$\begin{bmatrix} \frac{\partial x_1}{\partial \mathbf{v}} & \frac{\partial x_2}{\partial \mathbf{v}} & \dots & \frac{\partial x_n}{\partial \mathbf{v}} \end{bmatrix} \begin{bmatrix} a_{11} & a_{21} & \dots & a_{m1} \\ a_{12} & a_{22} & \dots & a_{m2} \\ \vdots & \vdots & \ddots & \vdots \\ a_{1n} & a_{2n} & \dots & a_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \tag{126}$$

$$= \frac{\partial \mathbf{x}}{\partial \mathbf{v}} \mathbf{A}\mathbf{x} + \frac{\partial \mathbf{x}}{\partial \mathbf{v}} \mathbf{A}^\top \mathbf{x} \tag{127}$$

$$\boxed{\frac{\partial \mathbf{x}^\top \mathbf{A}\mathbf{x}}{\partial \mathbf{v}} = \frac{\partial \mathbf{x}}{\partial \mathbf{v}} \left(\mathbf{A} + \mathbf{A}^\top\right) \mathbf{x} \in \mathbb{C}^m} \tag{128}$$

For the special case where $\mathbf{A}$ is symmetric, we get

$$\boxed{\frac{\partial \mathbf{x}^\top \mathbf{A}\mathbf{x}}{\partial \mathbf{v}} = 2\frac{\partial \mathbf{x}}{\partial \mathbf{v}} \mathbf{A}\mathbf{x} \in \mathbb{C}^m} \tag{129}$$

Note that the solution is much easier if we maintain the Matrix Calculus notation and apply the chain rule, that is,

$$\frac{\partial \mathbf{x}^\top \mathbf{A}\mathbf{x}}{\partial \mathbf{v}} = \frac{\partial \mathbf{x}}{\partial \mathbf{v}} \frac{\partial \mathbf{x}^\top \mathbf{A}\mathbf{x}}{\partial \mathbf{x}} = \frac{\partial \mathbf{x}}{\partial \mathbf{v}} \left(\mathbf{A}^\top + \mathbf{A}\right) \mathbf{x}, \tag{130}$$

where the last equality comes from (117).

5.14 $\frac{\partial \mathbf{b}^\top \mathbf{A}\mathbf{x}}{\partial \mathbf{x}} = \mathbf{A}^\top \mathbf{b}$

Let $\mathbf{x} \in \mathbb{R}^n$, $\mathbf{b} \in \mathbb{R}^m$ and $\mathbf{A} \in \mathbb{R}^{m \times n}$, where neither $\mathbf{b}$ nor $\mathbf{A}$ depend on $\mathbf{x}$. It follows that

$$\frac{\partial \mathbf{b}^\top \mathbf{A}\mathbf{x}}{\partial \mathbf{x}} = \frac{\partial}{\partial \mathbf{x}} \left(\begin{bmatrix} b_1 & b_2 & \dots & b_m \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \right) \tag{131}$$

$$= \frac{\partial}{\partial \mathbf{x}} \left(\begin{bmatrix} \sum_{i=1}^{m} a_{i1} b_i & \sum_{i=1}^{m} a_{i2} b_i & \dots & \sum_{i=1}^{m} a_{in} b_i \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \right) \tag{132}$$

$$= \frac{\partial}{\partial \mathbf{x}} \left(\sum_{i=1}^{m} \sum_{j=1}^{n} a_{ij} b_i x_j \right) \tag{133}$$

$$= \sum_{i=1}^{m} \sum_{j=1}^{n} a_{ij} b_i \frac{\partial x_j}{\partial \mathbf{x}} \tag{134}$$

$$= \begin{bmatrix} \sum_{i=1}^{m} a_{i1} b_i \\ \sum_{i=1}^{m} a_{i2} b_i \\ \vdots \\ \sum_{i=1}^{m} a_{in} b_i \end{bmatrix} \tag{135}$$

$$= \begin{bmatrix} a_{11} & a_{21} & \dots & a_{m1} \\ a_{12} & a_{22} & \dots & a_{m2} \\ \vdots & \vdots & \ddots & \vdots \\ a_{1n} & a_{2n} & \dots & a_{mn} \end{bmatrix} \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix} \tag{136}$$

$$\boxed{\frac{\partial \mathbf{b}^\top \mathbf{A}\mathbf{x}}{\partial \mathbf{x}} = \mathbf{A}^\top \mathbf{b} \in \mathbb{R}^n} \tag{137}$$

Note that this solution could solve by simply observing that $\mathbf{b}^\top \mathbf{A}$ is actually a linear transformation from $\mathbb{R}^n$ to $\mathbb{R}$. Thus,

$$\frac{\partial \mathbf{b}^\top \mathbf{A}\mathbf{x}}{\partial \mathbf{x}} = \left(\mathbf{b}^\top \mathbf{A}\right)^\top = \mathbf{A}^\top \mathbf{b}, \tag{138}$$

where the first equality comes from the (63).

5.15 $\frac{\partial \mathbf{x}^\top \mathbf{A}\mathbf{b}}{\partial \mathbf{x}} = \mathbf{A}\mathbf{b}$

Let $\mathbf{x} \in \mathbb{R}^m$, $\mathbf{b} \in \mathbb{R}^n$ and $\mathbf{A} \in \mathbb{R}^{m\times n}$, where neither $\mathbf{b}$ nor $\mathbf{A}$ depend on $\mathbf{x}$. The quickest way to solve it is to note that $\mathbf{x}^\top \mathbf{A}\mathbf{b} = \mathbf{b}^\top \mathbf{A}^\top \mathbf{x}$, which is the problem solved by the Section 5.14. Thus,

$$\frac{\partial \mathbf{x}^\top \mathbf{A}\mathbf{b}}{\partial \mathbf{x}} = \frac{\partial \mathbf{b}^\top \mathbf{A}^\top \mathbf{x}}{\partial \mathbf{x}} = \left(\mathbf{b}^\top \mathbf{A}^\top\right)^\top = \mathbf{A}\mathbf{b}, \tag{139}$$

where the second equality comes from (63). Nevertheless, here is the step-by-step

$$\frac{\partial \mathbf{x}^\top \mathbf{A}\mathbf{b}}{\partial \mathbf{x}} = \frac{\partial}{\partial \mathbf{x}} \left(\begin{bmatrix} x_1 & x_2 & \dots & x_m \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix} \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix} \right) \tag{140}$$

$$= \frac{\partial}{\partial \mathbf{x}} \left(\begin{bmatrix} \sum_{i=1}^{m} a_{i1} x_i & \sum_{i=1}^{m} a_{i2} x_i & \dots & \sum_{i=1}^{m} a_{in} x_i \end{bmatrix} \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix} \right) \tag{141}$$

$$= \frac{\partial}{\partial \mathbf{x}} \left(\sum_{i=1}^{m} \sum_{j=1}^{n} a_{ij} b_j x_i \right) \tag{142}$$

$$= \sum_{i=1}^{m} \sum_{j=1}^{n} a_{ij} b_j \frac{\partial x_i}{\partial \mathbf{x}} \tag{143}$$

$$= \begin{bmatrix} \sum_{j=1}^{n} a_{1j} b_j \\ \sum_{j=1}^{n} a_{2j} b_j \\ \vdots \\ \sum_{j=1}^{n} a_{nj} b_j \end{bmatrix} \tag{144}$$

$$= \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix} \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix} \tag{145}$$

$$\boxed{\frac{\partial \mathbf{x}^\top \mathbf{A}\mathbf{b}}{\partial \mathbf{x}} = \mathbf{A}\mathbf{b} \in \mathbb{R}^m} \tag{146}$$

5.16 $\dfrac{\partial \mathbf{y}^\top \mathbf{A}\mathbf{x}}{\partial \mathbf{v}} = \dfrac{\partial \mathbf{x}}{\partial \mathbf{v}} \mathbf{A}^\top \mathbf{y} + \dfrac{\partial \mathbf{y}}{\partial \mathbf{v}} \mathbf{A}\mathbf{x}$

Let $\mathbf{x} \in \mathbb{R}^n$, $\mathbf{y} \in \mathbb{R}^m$, $\mathbf{v} \in \mathbb{R}^p$, and $\mathbf{A} \in \mathbb{R}^{m \times n}$, where $\mathbf{x}$ and $\mathbf{y}$ depend on $\mathbf{v}$, but $\mathbf{A}$ does not. Therefore,

$$\frac{\partial \mathbf{y}^\top \mathbf{A}\mathbf{x}}{\partial \mathbf{v}} = \frac{\partial}{\partial \mathbf{v}} \left(\begin{bmatrix} y_1 & y_2 & \dots & y_m \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \right) \tag{147}$$

$$= \frac{\partial}{\partial \mathbf{v}} \left(\begin{bmatrix} \sum_{i=1}^{m} a_{i1} y_i & \sum_{i=1}^{m} a_{i2} y_i & \dots & \sum_{i=1}^{m} a_{in} y_i \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \right) \tag{148}$$

$$= \frac{\partial}{\partial \mathbf{v}} \left(\sum_{i=1}^{m} \sum_{j=1}^{n} a_{ij} x_j y_i \right) \tag{149}$$

$$= \sum_{i=1}^{m} \sum_{j=1}^{n} a_{ij} \frac{\partial y_i x_j}{\partial \mathbf{v}} \tag{150}$$

Recalling that $(fg)' = f'g + g'f$, we have that

$$
\begin{aligned}
\frac{\partial \mathbf{y}^\top \mathbf{A}\mathbf{x}}{\partial \mathbf{v}} &= \sum_{i=1}^{m}\sum_{j=1}^{n} a_{ij} y_i \frac{\partial x_j}{\partial \mathbf{v}} + \sum_{i=1}^{m}\sum_{j=1}^{n} a_{ij} x_j \frac{\partial y_i}{\partial \mathbf{v}} &(151)\\
&= \begin{bmatrix} \frac{\partial x_1}{\partial \mathbf{v}} & \frac{\partial x_2}{\partial \mathbf{v}} & \dots & \frac{\partial x_n}{\partial \mathbf{v}} \end{bmatrix} \begin{bmatrix} a_{11} & a_{21} & \dots & a_{m1} \\ a_{12} & a_{22} & \dots & a_{m2} \\ \vdots & \vdots & \ddots & \vdots \\ a_{1n} & a_{2n} & \dots & a_{mn} \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{bmatrix} + &(152)\\
&\quad \begin{bmatrix} \frac{\partial y_1}{\partial \mathbf{v}} & \frac{\partial y_2}{\partial \mathbf{v}} & \dots & \frac{\partial y_m}{\partial \mathbf{v}} \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} &(153)
\end{aligned}
$$

$$
\boxed{\frac{\partial \mathbf{y}^\top \mathbf{A}\mathbf{x}}{\partial \mathbf{v}} = \frac{\partial \mathbf{x}}{\partial \mathbf{v}} \mathbf{A}^\top \mathbf{y} + \frac{\partial \mathbf{y}}{\partial \mathbf{v}} \mathbf{A}\mathbf{x}} \tag{154}
$$

Even though this problem is trickier, we can find the same solution in a clever way by preserving the Matrix Calculus notation and applying the chain rule. Note that $\mathbf{y}^\top \mathbf{A}\mathbf{x}$ depends on both $\mathbf{x}$ and $\mathbf{y}$ which, in turn, depend on $\mathbf{v}$. Therefore (cf. (47)),

$$
\begin{aligned}
\frac{\partial \mathbf{y}^\top \mathbf{A}\mathbf{x}}{\partial \mathbf{v}} &= \frac{\partial \mathbf{x}}{\partial \mathbf{v}} \frac{\partial \mathbf{y}^\top \mathbf{A}\mathbf{x}}{\partial \mathbf{x}} + \frac{\partial \mathbf{y}}{\partial \mathbf{v}} \frac{\partial \mathbf{y}^\top \mathbf{A}\mathbf{x}}{\partial \mathbf{y}} &(155)\\
&= \frac{\partial \mathbf{x}}{\partial \mathbf{v}} \mathbf{A}^\top \mathbf{y} + \frac{\partial \mathbf{y}}{\partial \mathbf{v}} \mathbf{A}\mathbf{x}. &(156)
\end{aligned}
$$

5.17 $\frac{\partial \operatorname{tr}(\mathbf{A}\mathbf{X})}{\partial \mathbf{X}} = \frac{\partial \operatorname{tr}(\mathbf{X}\mathbf{A})}{\partial \mathbf{X}} = \mathbf{A}^\top$

Let $\mathbf{A} \in \mathbb{R}^{n\times m}$ and $\mathbf{X} \in \mathbb{R}^{m\times n}$, where $\mathbf{A}$ does not depend on the elements in $\mathbf{X}$.

$$
\begin{aligned}
\frac{\partial \operatorname{tr}(\mathbf{A}\mathbf{X})}{\partial \mathbf{X}} &= \frac{\partial}{\partial \mathbf{X}} \left(\operatorname{tr}\left(\begin{bmatrix} a_{11} & a_{12} & \dots & a_{1m} \\ a_{21} & a_{22} & \dots & a_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nm} \end{bmatrix} \begin{bmatrix} x_{11} & x_{12} & \dots & x_{1n} \\ x_{21} & x_{22} & \dots & x_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ x_{m1} & x_{m2} & \dots & x_{mn} \end{bmatrix} \right) \right) \\
&= \frac{\partial}{\partial \mathbf{X}} \left(\sum_{i=1}^{n}\sum_{j=1}^{m} a_{ij} x_{ji} \right) \\
&= \begin{bmatrix}
\frac{\partial}{\partial x_{11}} \left(\sum_{i=1}^{n}\sum_{j=1}^{m} a_{ij} x_{ji} \right) & \frac{\partial}{\partial x_{12}} \left(\sum_{i=1}^{n}\sum_{j=1}^{m} a_{ij} x_{ji} \right) & \dots & \frac{\partial}{\partial x_{1n}} \left(\sum_{i=1}^{n}\sum_{j=1}^{m} a_{ij} x_{ji} \right) \\
\frac{\partial}{\partial x_{21}} \left(\sum_{i=1}^{n}\sum_{j=1}^{m} a_{ij} x_{ji} \right) & \frac{\partial}{\partial x_{22}} \left(\sum_{i=1}^{n}\sum_{j=1}^{m} a_{ij} x_{ji} \right) & \dots & \frac{\partial}{\partial x_{2n}} \left(\sum_{i=1}^{n}\sum_{j=1}^{m} a_{ij} x_{ji} \right) \\
\vdots & \vdots & \ddots & \vdots \\
\frac{\partial}{\partial x_{m1}} \left(\sum_{i=1}^{n}\sum_{j=1}^{m} a_{ij} x_{ji} \right) & \frac{\partial}{\partial x_{m2}} \left(\sum_{i=1}^{n}\sum_{j=1}^{m} a_{ij} x_{ji} \right) & \dots & \frac{\partial}{\partial x_{mn}} \left(\sum_{i=1}^{n}\sum_{j=1}^{m} a_{ij} x_{ji} \right)
\end{bmatrix} \\
&= \begin{bmatrix} a_{11} & a_{21} & \dots & a_{n1} \\ a_{12} & a_{22} & \dots & a_{n2} \\ \vdots & \vdots & \ddots & \vdots \\ a_{1m} & a_{2m} & \dots & a_{mn} \end{bmatrix} = \mathbf{A}^\top
\end{aligned}
$$

Finally, recalling that $\operatorname{tr}(\mathbf{AB}) = \operatorname{tr}(\mathbf{BA})$, we have that (note that $\mathbf{AB}$ and $\mathbf{BA}$ must be square)

$$\boxed{\frac{\partial \operatorname{tr}(\mathbf{AX})}{\partial \mathbf{X}} = \frac{\partial \operatorname{tr}(\mathbf{XA})}{\partial \mathbf{X}} = \mathbf{A}^\top} \tag{157}$$

I have no idea how to make this solution simpler.

5.18 $\frac{\partial \operatorname{tr}\left(\mathbf{XAX}^\top\right)}{\partial \mathbf{X}} = 2\mathbf{XA}$

Let $\mathbf{A} \in \mathbb{S}^n$ be a symmetric matrix ($\mathbb{S}^n$ denotes the subspace of the symmetric matrices in $\mathbb{R}^{n\times n}$) and $\mathbf{X} \in \mathbb{R}^{n\times n}$, where $\mathbf{A}$ does not depend on the elements in $\mathbf{X}$.

$$\begin{aligned}
\frac{\partial \operatorname{tr}\left(\mathbf{XAX}^\top\right)}{\partial \mathbf{X}} &= \frac{\partial}{\partial \mathbf{X}}\left(\operatorname{tr}\left(\begin{bmatrix} x_{11} & x_{12} & \dots & x_{1n} \\ x_{21} & x_{22} & \dots & x_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ x_{n1} & x_{n2} & \dots & x_{nn} \end{bmatrix}\begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{bmatrix}\begin{bmatrix} x_{11} & x_{21} & \dots & x_{n1} \\ x_{12} & x_{22} & \dots & x_{n2} \\ \vdots & \vdots & \ddots & \vdots \\ x_{1n} & x_{2n} & \dots & x_{nn} \end{bmatrix}\right)\right) \\
&= \frac{\partial}{\partial \mathbf{X}}\left[\sum_{i=1}^{n}\sum_{k=1}^{n}\sum_{j=1}^{n} x_{ik}a_{kj}x_{ij}\right] \\
&= \begin{bmatrix}
\frac{\partial}{\partial x_{11}}\left(\sum_{i=1}^{n}\sum_{k=1}^{n}\sum_{j=1}^{n} x_{ik}a_{kj}x_{ij}\right) & \frac{\partial}{\partial x_{12}}\left(\sum_{i=1}^{n}\sum_{k=1}^{n}\sum_{j=1}^{n} x_{ik}a_{kj}x_{ij}\right) & \dots & \frac{\partial}{\partial x_{1n}}\left(\sum_{i=1}^{n}\sum_{k=1}^{n}\sum_{j=1}^{n} x_{ik}a_{kj}x_{ij}\right) \\
\frac{\partial}{\partial x_{21}}\left(\sum_{i=1}^{n}\sum_{k=1}^{n}\sum_{j=1}^{n} x_{ik}a_{kj}x_{ij}\right) & \frac{\partial}{\partial x_{22}}\left(\sum_{i=1}^{n}\sum_{k=1}^{n}\sum_{j=1}^{n} x_{ik}a_{kj}x_{ij}\right) & \dots & \frac{\partial}{\partial x_{2n}}\left(\sum_{i=1}^{n}\sum_{k=1}^{n}\sum_{j=1}^{n} x_{ik}a_{kj}x_{ij}\right) \\
\vdots & \vdots & \ddots & \vdots \\
\frac{\partial}{\partial x_{n1}}\left(\sum_{i=1}^{n}\sum_{k=1}^{n}\sum_{j=1}^{n} x_{ik}a_{kj}x_{ij}\right) & \frac{\partial}{\partial x_{n2}}\left(\sum_{i=1}^{n}\sum_{k=1}^{n}\sum_{j=1}^{n} x_{ik}a_{kj}x_{ij}\right) & \dots & \frac{\partial}{\partial x_{nn}}\left(\sum_{i=1}^{n}\sum_{k=1}^{n}\sum_{j=1}^{n} x_{ik}a_{kj}x_{ij}\right)
\end{bmatrix} \\
&= \begin{bmatrix}
\sum_{j=1}^n a_{1j}x_{1j} + \sum_{k=1}^n x_{1k}a_{k1} & \sum_{j=1}^n a_{2j}x_{1j} + \sum_{k=1}^n x_{1k}a_{k2} & \dots & \sum_{j=1}^n a_{nj}x_{1j} + \sum_{k=1}^n x_{1k}a_{kn} \\
\sum_{j=1}^n a_{1j}x_{2j} + \sum_{k=1}^n x_{2k}a_{k1} & \sum_{j=1}^n a_{2j}x_{2j} + \sum_{k=1}^n x_{2k}a_{k2} & \dots & \sum_{j=1}^n a_{nj}x_{2j} + \sum_{k=1}^n x_{2k}a_{kn} \\
\vdots & \vdots & \ddots & \vdots \\
\sum_{j=1}^n a_{1j}x_{nj} + \sum_{k=1}^n x_{nk}a_{k1} & \sum_{j=1}^n a_{2j}x_{nj} + \sum_{k=1}^n x_{nk}a_{k2} & \dots & \sum_{j=1}^n a_{nj}x_{nj} + \sum_{k=1}^n x_{nk}a_{kn}
\end{bmatrix}
\end{aligned}$$

Since $a_{ij} = a_{ji}$, it follows that

$$\frac{\partial \operatorname{tr}\left(\mathbf{XAX}^\top\right)}{\partial \mathbf{X}} = \begin{bmatrix}
2\sum_{k=1}^n x_{1k}a_{k1} & 2\sum_{k=1}^n x_{1k}a_{k2} & \dots & 2\sum_{k=1}^n x_{1k}a_{kn} \\
2\sum_{k=1}^n x_{2k}a_{k1} & 2\sum_{k=1}^n x_{2k}a_{k2} & \dots & 2\sum_{k=1}^n x_{2k}a_{kn} \\
\vdots & \vdots & \ddots & \vdots \\
2\sum_{k=1}^n x_{nk}a_{k1} & 2\sum_{k=1}^n x_{nk}a_{k2} & \dots & 2\sum_{k=1}^n x_{nk}a_{kn}
\end{bmatrix} = 2\mathbf{XA} \tag{158}$$

Therefore,

$$\boxed{\frac{\partial \operatorname{tr}\left(\mathbf{XAX}^\top\right)}{\partial \mathbf{X}} = 2\mathbf{XA}} \tag{159}$$

I have no idea how to make this solution simpler.

5.19 $\frac{\partial|\mathbf{X}|}{\partial\mathbf{X}} = \mathrm{adj}\,(\mathbf{X})$

Let $\mathbf{X} \in \mathbb{R}^{n\times n}$. Through Laplace expansion (cofactor expansion), we can rewrite the determinant of $\mathbf{X}$ as the sum of the cofactors of any row or column, multiplied by its generating element, that is

$$|\mathbf{X}| = \sum_{i=1}^{n} x_{ki}|\mathbf{C}_{ki}| = \sum_{i=1}^{n} x_{ik}|\mathbf{C}_{ik}| \quad \forall\ k \in \{1, 2, ..., n\}\,, \tag{160}$$

where $\mathbf{C}_{ij}$ denotes the cofactor matrix of $\mathbf{X}$ generated from element x_{ij}. It is worth noting that the cofactor of $\mathbf{C}_{ij}$ is independent of the value of any element (i, j) in $\mathbf{X}$. Therefore, it follows that

$$\frac{\partial|\mathbf{X}|}{\partial\mathbf{X}} = \frac{\partial}{\partial\mathbf{X}}\left(\sum_{i=1}^{n} x_{ki}|\mathbf{C}_{ki}|\right) \quad \forall\ k \in \{1, 2, ..., n\} \tag{161}$$

$$= \frac{\partial}{\partial\mathbf{X}}\left(\begin{bmatrix} \sum_{i=1}^{n} x_{1i}|\mathbf{C}_{1i}| & \sum_{i=1}^{n} x_{1i}|\mathbf{C}_{1i}| & \dots & \sum_{i=1}^{n} x_{1i}|\mathbf{C}_{1i}| \\ \sum_{i=1}^{n} x_{2i}|\mathbf{C}_{2i}| & \sum_{i=1}^{n} x_{2i}|\mathbf{C}_{2i}| & \dots & \sum_{i=1}^{n} x_{2i}|\mathbf{C}_{2i}| \\ \vdots & \vdots & \ddots & \vdots \\ \sum_{i=1}^{n} x_{ni}|\mathbf{C}_{ni}| & \sum_{i=1}^{n} x_{ni}|\mathbf{C}_{ni}| & \dots & \sum_{i=1}^{n} x_{ni}|\mathbf{C}_{ni}| \end{bmatrix}\right) \tag{162}$$

$$= \begin{bmatrix} \frac{\partial}{\partial x_{11}}\left(\sum_{i=1}^{n} x_{1i}|\mathbf{C}_{1i}|\right) & \frac{\partial}{\partial x_{12}}\left(\sum_{i=1}^{n} x_{1i}|\mathbf{C}_{1i}|\right) & \dots & \frac{\partial}{\partial x_{13}}\left(\sum_{i=1}^{n} x_{1i}|\mathbf{C}_{1i}|\right) \\ \frac{\partial}{\partial x_{21}}\left(\sum_{i=1}^{n} x_{2i}|\mathbf{C}_{2i}|\right) & \frac{\partial}{\partial x_{22}}\left(\sum_{i=1}^{n} x_{2i}|\mathbf{C}_{2i}|\right) & \dots & \frac{\partial}{\partial x_{33}}\left(\sum_{i=1}^{n} x_{2i}|\mathbf{C}_{2i}|\right) \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial}{\partial x_{n1}}\left(\sum_{i=1}^{n} x_{ni}|\mathbf{C}_{ni}|\right) & \frac{\partial}{\partial x_{n2}}\left(\sum_{i=1}^{n} x_{ni}|\mathbf{C}_{ni}|\right) & \dots & \frac{\partial}{\partial x_{n3}}\left(\sum_{i=1}^{n} x_{ni}|\mathbf{C}_{ni}|\right) \end{bmatrix} \tag{163}$$

$$= \begin{bmatrix} |\mathbf{C}_{11}| & |\mathbf{C}_{12}| & \dots & |\mathbf{C}_{1n}| \\ |\mathbf{C}_{21}| & |\mathbf{C}_{22}| & \dots & |\mathbf{C}_{2n}| \\ \vdots & \vdots & \ddots & \vdots \\ |\mathbf{C}_{n1}| & |\mathbf{C}_{n2}| & \dots & |\mathbf{C}_{nn}| \end{bmatrix} \tag{164}$$

$$\boxed{\frac{\partial|\mathbf{X}|}{\partial\mathbf{X}} = \mathrm{adj}\,(\mathbf{X})} \tag{165}$$

I have no idea how to make this solution simpler.

5.20 $\frac{\partial\mathbf{A}^{-1}}{\partial\alpha} = -\left(\mathbf{A}^{-1}\right)^{\top}\frac{\partial\mathbf{A}}{\partial\alpha}\left(\mathbf{A}^{-1}\right)^{\top}$

Let $\mathbf{A} \in \mathbb{R}^{m\times n}$ and $\alpha \in \mathbb{R}$. Remember that $\mathbf{A}^{-1}\mathbf{A} = \mathbf{I}$. Differentiating both sides of this equation with respect to α, we get

$$\frac{\partial\mathbf{A}^{-1}\mathbf{A}}{\partial\alpha} = \frac{\partial\mathbf{I}}{\partial\alpha} = \mathbf{0}_{m\times n}, \tag{166}$$

where $\mathbf{0}_{m\times n}$ is a zero matrix with dimension $m \times n$. By applying the product rule of a matrix-matrix derivate, we get (cf. (57))

$$\frac{\partial\mathbf{A}^{-1}\mathbf{A}}{\partial\alpha} = \frac{\partial\mathbf{A}}{\partial\alpha}\left(\mathbf{A}^{-1}\right)^{\top} + \mathbf{A}^{\top}\frac{\partial\mathbf{A}^{-1}}{\partial\alpha} = \mathbf{0}_{m\times n} \tag{167}$$

Using the property $\left(\mathbf{A}^\top\right)^{-1} = \left(\mathbf{A}^{-1}\right)^\top$ and rearranging this expression, we get

$$\boxed{\frac{\partial \mathbf{A}^{-1}}{\partial \alpha} = -\left(\mathbf{A}^{-1}\right)^\top \frac{\partial \mathbf{A}}{\partial \alpha} \left(\mathbf{A}^{-1}\right)^\top} \tag{168}$$

References

[1] Phoebus J Dhrymes and Phoebus J Dhrymes. *Mathematics for econometrics*. Vol. 984. Springer, 1978.

[2] Randal J Barnes. "Matrix differentiation". In: *Springs Journal* (2006), pp. 1–9.

[3] Aarti Singh. *Lecture notes in Introduction to Machine Learning*. 2013-2016. URL: https://www.cs.cmu.edu/~aarti/Class/10315_Spring22/315S22_Rec4.pdf.

[4] *Matrix calculus - Wikipedia*. (Accessed on 09/22/2022). URL: https://en.wikipedia.org/wiki/Matrix_calculus#Numerator-layout_notation.

[5] Shayle R Searle and Andre I Khuri. *Matrix algebra useful for statistics*. John Wiley & Sons, 2017.

[6] Are Hjørungnes. *Complex-valued matrix derivatives: with applications in signal processing and communications*. Cambridge University Press, 2011.

[7] Are Hjorungnes and David Gesbert. "Complex-valued matrix differentiation: Techniques and key results". In: *IEEE Transactions on Signal Processing* 55.6 (2007), pp. 2740–2746.

[8] Jon Dattorro. *Convex Optimization & Euclidean Distance Geometry*. Lulu. com, 2010. ISBN: 0-615-19368-4.

[9] Christopher M Bishop and Nasser M Nasrabadi. *Pattern Recognition and Machine Learning*. Vol. 4. 4. Springer, 2006.

[10] Simon S Haykin et al. *Neural networks and learning machines*. 2009.

[11] Kaare Brandt Petersen, Michael Syskind Pedersen, et al. "The matrix cookbook". In: *Technical University of Denmark* 7.15 (2008), p. 510.

[12] Marc Peter Deisenroth, A. Aldo Faisal, and Cheng Soon Ong. *Mathematics for Machine Learning*. Cambridge University Press, 2020. URL: https://books.google.com.br/books?hl=en&lr=&id=pFjPDwAAQBAJ&oi=fnd&pg=PR9&dq=Mathematics+for+Machine+Learning&ots=VMgn3EG3Db&sig=Ffu2Gxr5OnHGzZKRWb1mrAIkhRQ (visited on 03/22/2024).

[13] Old and New Matrix Algebra Useful for Statistics. *Lecture notes of Macroeconomics*. 2000. URL: https://tminka.github.io/papers/matrix/minka-matrix.pdf.

[14] Terence Parr and Jeremy Howard. *The Matrix Calculus You Need For Deep Learning*. July 2, 2018. arXiv: 1802.01528 [cs, stat]. URL: http://arxiv.org/abs/1802.01528 (visited on 05/05/2023). preprint.

[15] James Stewart, Daniel K Clegg, and Saleem Watson. *Calculus: early transcendentals*. Cengage Learning, 2020.

[16] Paulo SR Diniz et al. *Adaptive filtering*. Vol. 4. Springer, 1997.

[17] Simon S Haykin. *Adaptive Filter Theory*. Pearson Education India, 2002. ISBN: 81-317-0869-1.

[18] J. E. Dennis and Robert B. Schnabel. *Numerical Methods for Unconstrained Optimization and Nonlinear Equations*. Society for Industrial and Applied Mathematics, Jan. 1996. ISBN: 978-0-89871-364-0 978-1-61197-120-0. DOI: 10.1137/1.9781611971200. URL: http://epubs.siam.org/doi/book/10.1137/1.9781611971200 (visited on 03/22/2024).

[19] Simon Haykin. *Neural Networks and Learning Machines, 3/E*. Pearson Education India, 2009. ISBN: 93-325-8625-X.

[20] Rubem V. Pacelli et al. "A Data-Based Estimation of Power-Law Coefficients for Rainfall via Levenberg-Marquardt Algorithm: Results from the West African Testbed". In: *Proceedings of the 6th International Symposium on Uncertainty Quantification and Stochastic Modelling*. Ed. by José Eduardo Souza De Cursi. Cham: Springer International Publishing, 2024, pp. 190–201. ISBN: 978-3-031-47035-6 978-3-031-47036-3. DOI: 10.1007/978-3-031-47036-3_17. URL: https://link.springer.com/10.1007/978-3-031-47036-3_17 (visited on 03/22/2024).

[21] Kaj Madsen, Hans Nielsen, and O Tingleff. “Methods for Non-Linear Least Squares Problems (2nd Ed.)” In: (Jan. 1, 2004), p. 60.

[22] Cesar Augusto Taconeli. *Lecture notes in Análise de Regressão Linear*. 2018. URL: https://docs.ufpr.br/~taconeli/CE07118/Algebra.pdf.

[23] Paul Klein. *Lecture notes of Macroeconomics*. 1999. URL: http://paulklein.ca/newsite/teaching/calcvec.pdf.

[24] *P.838 : Specific Attenuation Model for Rain for Use in Prediction Methods*. URL: https://www.itu.int/rec/R-REC-P.838 (visited on 02/16/2023).

[25] Jules Aarons. “Construction of a Model of Equatorial Scintillation Intensity”. In: *Radio Science* 20.3 (1985), pp. 397–402.

[26] Jules Aarons. “Global Morphology of Ionospheric Scintillations”. In: *Proceedings of the IEEE* 70.4 (1982), pp. 360–378. ISSN: 0018-9219.

[27] MA Abdu, JA Bittencourt, and IS Batista. “Magnetic Declination Control of the Equatorial F Region Dynamo Electric Field Development and Spread F”. In: *Journal of Geophysical Research: Space Physics* 86.A13 (1981), pp. 11443–11446. ISSN: 0148-0227.

[28] MA Abdu. “Outstanding Problems in the Equatorial Ionosphere–Thermosphere Electrodynamics Relevant to Spread F”. In: *Journal of Atmospheric and Solar-Terrestrial Physics* 63.9 (2001), pp. 869–884. ISSN: 1364-6826.

[29] Sílvio A Abrantes. “Introdução à Sincronização Em Modulações Digitais”. In: *Março de* (2007).

[30] Sílvio A Abrantes. “Recuperação Digital Da Temporização Com Amostragem Assíncrona–Parte 1: Transmissão Em Banda-Base”. In: (2010).

[31] E. L. Afraimovich et al. “Characteristics of Small-Scale Ionospheric Irregularities as Deduced from Scintillation Observations of Radio Signals from Satellites ETS-2 and Polar Bear 4 at Irkutsk”. In: *Radio science* 29.04 (1994), pp. 839–855.

[32] D.M. Akos et al. “Direct Bandpass Sampling of Multiple Distinct RF Signals”. In: *IEEE Transactions on Communications* 47.7 (July 1999), pp. 983–988. ISSN: 00906778. DOI: 10.1109/26.774848. URL: http://ieeexplore.ieee.org/document/774848/ (visited on 01/21/2023).

[33] Jennifer Alvarez and Buddy Walls. “Constellations, Clusters, and Communication Technology: Expanding Small Satellite Access to Space”. In: 2016 IEEE Aerospace Conference. IEEE, 2016, pp. 1–11. ISBN: 1-4673-7676-0.

[34] TM Apostol. *Calculus, 2nd Edn., Vol. 2*. 1967.

[35] Tom M Apostol. *Calculus, Volume 1*. John Wiley & Sons, 1991. ISBN: 0-471-00005-1.

[36] Constantine A Balanis. *Advanced Engineering Electromagnetics*. John Wiley & Sons, 2012. ISBN: 0-470-58948-5.

[37] *Bandwidth-Efficient Modulations: Summary of Definition, Implementation, and Performance*. 2018.

[38] Santimay Basu et al. “250 MHz/GHz Scintillation Parameters in the Equatorial, Polar, and Auroral Environments”. In: *IEEE journal on selected areas in communications* 5.2 (1987), pp. 102–115. ISSN: 0733-8716.

[39] Sunanda Basu, Santimay Basu, and B. K. Khan. “Model of Equatorial Scintillations from In-Situ Measurements”. In: *Radio Science* 11.10 (1976), pp. 821–832.

[40] Su Basu, Sa Basu, and W. B. Hanson. “The Role of In-Situ Measurements in Scintillation Modelling”. In: *In: Symposium on the Effect of the Ionosphere on Radiowave Systems*. 1981.

[41] Yoshua Bengio, Aaron Courville, and Pascal Vincent. “Representation Learning: A Review and New Perspectives”. In: *IEEE transactions on pattern analysis and machine intelligence* 35.8 (2013), pp. 1798–1828.

[42] Yannick Beniguel. “Characterisation of Equatorial Scintillations For Transionospheric Links”. In: *Proceedings of the 13th International Technical Meeting of the Satellite Division of The Institute of Navigation (ION GPS 2000)*. 2000, pp. 688–695.

[43] Y. Béniguel. *GISM Technical Manual*. IEEA, Jan. 9, 2011, p. 21. URL: http://www.ieea.fr/help/gism-technical.pdf (visited on 05/19/2023).

[44] Y. Béniguel. *GISM User Manual*. 6.53. IEEA, Feb. 27, 2011, p. 33. URL: http://www.ieea.fr/help/gism-user-manual.pdf (visited on 05/19/2023).

[45] Yannick Béniguel and Pierrick Hamel. "A Global Ionosphere Scintillation Propagation Model for Equatorial Regions". In: *Journal of Space Weather and Space Climate* 1.1 (2011), A04.

[46] Y. Béniguel. "Global Ionospheric Propagation Model (GIM): A Propagation Model for Scintillations of Transmitted Signals". In: *Radio science* 37.3 (2002), pp. 1–14.

[47] Y. Béniguel and S. Buonomo. "A Multiple Phase Screen Ionospheric Propagation Model to Estimate the Fluctuations of Transmitted Signals". In: *Physics and Chemistry of the Earth, Part C: Solar, Terrestrial & Planetary Science* 24.4 (1999), pp. 333–338.

[48] Yannick Béniguel et al. "Scintillations Effects on Satellite to Earth Links for Telecommunication and Navigation Purposes". In: *Annals of geophysics* 47 (2-3 Sup. 2004). ISSN: 2037-416X.

[49] Arnold Berger. *Embedded Systems Design: An Introduction to Processes, Tools, and Techniques*. CRC Press, 2001. ISBN: 0-429-17936-7.

[50] Kiran Bhandarkar et al. "Design and Implementation of FPGA Based Reconfigurable Modulator for Satellite Applications". In: TENCON 2017-2017 IEEE Region 10 Conference. IEEE, 2017, pp. 1427–1432. ISBN: 1-5090-1134-X.

[51] Debopam Bhattacherjee et al. "Gearing up for the 21st Century Space Race". In: Proceedings of the 17th ACM Workshop on Hot Topics in Networks. 2018, pp. 113–119.

[52] Dieter Bilitza. "IRI the International Standard for the Ionosphere". In: *Advances in Radio Science* 16 (Sept. 4, 2018), pp. 1–11. ISSN: 1684-9973. DOI: 10.5194/ars-16-1-2018. URL: https://ars.copernicus.org/articles/16/1/2018/ (visited on 08/23/2023).

[53] K Bischoff and B Chytil. "A Note on Scintillation Indices". In: *Planetary and Space Science* 17.5 (1969), pp. 1059–1066. ISSN: 0032-0633.

[54] Henry George Booker, John Ashworth Ratcliffe, and D. H. Shinn. "Diffraction from an Irregular Screen with Applications to Ionospheric Problems". In: *Philosophical Transactions of the Royal Society of London. Series A, Mathematical and Physical Sciences* 242.856 (1950), pp. 579–607.

[55] S. A. Bowhill. "Statistics of a Radio Wave Diffracted by a Random Ionosphere". In: *J. Res. Nat. Bur. Stand., Sect. D* 65 (1961), pp. 275–292.

[56] Stephen Boyd and Lieven Vandenberghe. "Additional Exercises for Convex Optimization". In: ().

[57] Stephen Boyd, Stephen P. Boyd, and Lieven Vandenberghe. *Convex Optimization*. Cambridge university press, 2004.

[58] Stephen Boyd and Michael Grant. "Disciplined Convex Programming". In: *Convex optimization* (), p. 53.

[59] Gerd Brewka. *Artificial Intelligence—a Modern Approach by Stuart Russell and Peter Norvig, Prentice Hall. Series in Artificial Intelligence, Englewood Cliffs, NJ.* Vol. 11. Mar. 1996. URL: https://www.cambridge.org/core/product/identifier/S0269888900007724/type/journal_article (visited on 01/21/2023).

[60] B. H. Briggs and I. A. Parkin. "On the Variation of Radio Star and Satellite Scintillations with Zenith Angle". In: *Journal of Atmospheric and Terrestrial Physics* 25.6 (June 1, 1963), pp. 339–366. ISSN: 0021-9169. DOI: 10.1016/0021-9169(63)90150-8. URL: https://www.sciencedirect.com/science/article/pii/0021916963901508 (visited on 04/28/2023).

[61] Robert Grover Brown and Patrick YC Hwang. *Introduction to Random Signals and Applied Kalman Filtering: With MATLAB Exercises and Solutions*. 1997.

[62] K. G. Budden. "The Amplitude Fluctuations of the Radio Wave Scattered from a Thick Ionospheric Layer with Weak Irregularities". In: *Journal of Atmospheric and Terrestrial Physics* 27.2 (1965), pp. 155–172.

[63] Kenneth George Budden. *The Propagation of Radio Waves: The Theory of Radio Waves of Low Power in the Ionosphere and Magnetosphere*. Cambridge University Press, 1988.

[64] K. G. Budden. “The Theory of the Correlation of Amplitude Fluctuations of Radio Signals at Two Frequencies, Simultaneously Scattered by the Ionosphere”. In: *Journal of Atmospheric and Terrestrial Physics* 27.8 (1965), pp. 883–897.

[65] Adriano José Camps Carmona et al. “Ionospheric Scintillation Models: An Inter-Comparison Study Using GNSS Data”. In: (2023).

[66] Adriano Camps et al. “Improved Modelling of Ionospheric Disturbances for Remote Sensing and Navigation”. In: *2017 IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*. IEEE, 2017, pp. 2682–2685.

[67] Adriano Camps et al. “Ionospheric Scintillation Model Limitations and Impact in GNSS-R Missions”. In: *IGARSS 2020-2020 IEEE International Geoscience and Remote Sensing Symposium*. IEEE, 2020, pp. 5937–5940.

[68] Steven Chapra and Raymond Canale. *Numerical Methods for Engineers*. 7ª edição. New York, NY: McGraw Hill, Jan. 24, 2014. ISBN: 978-0-07-339792-4.

[69] Rama Chellappa and Sergios Theodoridis. *Signal Processing Theory and Machine Learning*. Academic Press, 2014. ISBN: 0-12-396502-0.

[70] David Keun Cheng. *Field and Wave Electromagnetics*. Pearson Education India, 1989.

[71] B. Chytil. “The Distribution of Amplitude Scintillation and the Conversion of Scintillation Indices”. In: *Journal of Atmospheric and Terrestrial Physics* 29.9 (Sept. 1, 1967), pp. 1175–1177. ISSN: 0021-9169. DOI: 10.1016/0021-9169(67)90151-1. URL: https://www.sciencedirect.com/science/article/pii/0021916967901511 (visited on 04/28/2023).

[72] Robert S Conker et al. “Modeling the Effects of Ionospheric Scintillation on GPS/Satellite-Based Augmentation System Availability”. In: *Radio Science* 38.1 (2003), pp. 1–1. ISSN: 1944-799X.

[73] Thomas H Cormen et al. *Introduction to Algorithms*. MIT press, 2022. ISBN: 0-262-36750-5.

[74] R.K. Crane. “Ionospheric Scintillation”. In: *Proceedings of the IEEE* 65.2 (Feb. 1977), pp. 180–199. ISSN: 1558-2256. DOI: 10.1109/PROC.1977.10456.

[75] *CubeSat Design Specification (CDS)*.

[76] Aldo N D’Andrea, Umberto Mengali, and Ruggero Reggiannini. “The Modified Cramer-Rao Bound and Its Application to Synchronization Problems”. In: *IEEE Transactions on Communications* 42.234 (1994), pp. 1391–1399. ISSN: 0090-6778.

[77] Aldo N D’Andrea, Umberto Mengali, and Michele Morelli. “Symbol Timing Estimation with CPM Modulation”. In: *IEEE Transactions on communications* 44.10 (1996), pp. 1362–1372. ISSN: 0090-6778.

[78] *The DCP Ruleset — CVX Users’ Guide*. URL: http://cvxr.com/cvx/doc/dcp.html (visited on 11/27/2022).

[79] Antonio Macilio Pereira de Lucena et al. “Design of a Fully Digital BPSK Demodulator Integrated into a TT&C Satellite Transponder”. In: *IEEE Latin America Transactions* 18.09 (2020), pp. 1511–1520. ISSN: 1548-0992.

[80] A. M. P. de Lucena, F. de A. T. F. da Silva, and A. S. da Silva. “Scintillation Effects in S-Band Telemetry Link of INPE’s Earth Station in Cuiaba-Brazil”. In: *Radioengineering* 30.4 (Sept. 15, 2021), pp. 739–748. ISSN: 1210-2512. DOI: 10.13164/re.2021.0739. URL: https://www.radioeng.cz/fulltexts/2021/21_04_0739_0748.pdf (visited on 01/24/2023).

[81] ER De Paula et al. “Equatorial Anomaly Effects on GPS Scintillations in Brazil”. In: *Advances in Space Research* 31.3 (2003), pp. 749–754. ISSN: 0273-1177.

[82] Paulo SR Diniz. *Adaptive Filtering: Algorithms and Practical Implementation*. Nowell, MA: Kluwer Academic Publishers, 2002.

[83] Paulo SR Diniz, Eduardo AB Da Silva, and Sergio L Netto. *Digital Signal Processing: System Analysis and Design*. Cambridge University Press, 2010. ISBN: 1-139-49157-1.

[84] Daniel Egea-Roca et al. “GNSS User Technology: State-of-the-Art and Future Trends”. In: *IEEE Access* 10 (2022), pp. 39939–39968. ISSN: 2169-3536.

[85] Matteo Emanuellia et al. “Nanosatellites and Their Demand for Changes in Space Policy”. In: 67th International Astronautical Congress (IAC). 2016.

[86] Susanna S Epp. *Discrete Mathematics with Applications*. Cengage learning, 2010. ISBN: 1-133-16866-3.

[87] Cecil William Farrow. “Continuously Variable Digital Delay Circuit”. Pat. CA1298354C (CA). American Telephone and Telegraph Co Inc. Mar. 31, 1992. URL: https://patents.google.com/patent/CA1298354C/en (visited on 01/19/2023).

[88] Carles Fernandez. “Adaptive GNSS Carrier Tracking Under Ionospheric Scintillation: Estimation vs. Mitigation”. In: *IEEE Communications Letters* (Jan. 1, 2015). URL: https://www.academia.edu/87064588/Adaptive_GNSS_Carrier_Tracking_Under_Ionospheric_Scintillation_Estimation_vs_Mitigation (visited on 02/13/2024).

[89] M Gardner Floyd. “Interpolation in Digital Modems–Part I: Fundamentals”. In: *IEEE Transactions on Communications* 41.3 (1993), pp. 501–507.

[90] Friederike Fohlmeister, Felix Antreich, and Josef A. Nossek. “Dual Kalman Filtering Based GNSS Phase Tracking for Scintillation Mitigation”. In: *2018 IEEE/ION Position, Location and Navigation Symposium (PLANS)*. 2018 IEEE/ION Position, Location and Navigation Symposium (PLANS). Apr. 2018, pp. 1151–1158. DOI: 10.1109/PLANS.2018.8373499.

[91] Friederike Fohlmeister. “GNSS Carrier Phase Tracking under Ionospheric Scintillations”. thesis. Technische Universität München, Apr. 2021. 118 pp. URL: https://elib.dlr.de/185344/ (visited on 02/28/2023).

[92] Biagio Forte and Sandro M. Radicella. “Comparison of Ionospheric Scintillation Models with Experimental Data for Satellite Navigation Applications”. In: *Annals of Geophysics* (2005).

[93] E. J. Fremouw and C. L. Rino. “An Empirical Model for Average F-layer Scintillation at VHF/UHF”. In: *Radio science* 8.3 (1973), pp. 213–222.

[94] E. J. Fremouw and J. A. Secan. “Modeling and Scientific Application of Scintillation Results”. In: *Radio Science* 19.03 (1984), pp. 687–694.

[95] E. J. Fremouw, R. C. Livingston, and Deborah A. Miller. “On the Statistics of Scintillating Signals”. In: *Journal of Atmospheric and Terrestrial Physics* 42.8 (Aug. 1, 1980), pp. 717–731. ISSN: 0021-9169. DOI: 10.1016/0021-9169(80)90055-0. URL: https://www.sciencedirect.com/science/article/pii/0021916980900550 (visited on 05/11/2023).

[96] *Functions*. URL: https://dcp.stanford.edu/functions (visited on 12/08/2022).

[97] Wanxuan Fu et al. “Real-Time Ionospheric Scintillation Monitoring”. In: Proceedings of the 12th International Technical Meeting of the Satellite Division of The Institute of Navigation (ION GPS 1999). 1999, pp. 1461–1472.

[98] Hélène Galiègue, Laurent Féral, and Vincent Fabbro. “Validity of 2-D Electromagnetic Approaches to Estimate Log-Amplitude and Phase Variances Due to 3-D Ionospheric Irregularities”. In: *Journal of Geophysical Research: Space Physics* 122.1 (2017), pp. 1410–1427.

[99] Artur Gaysin, Vladimir Fadeev, and Marko Hennhöfer. “Survey of Modulation and Coding Schemes for Application in CubeSat Systems”. In: 2017 Systems of Signal Synchronization, Generating and Processing in Telecommunications (SINKHROINFO). IEEE, 2017, pp. 1–7. ISBN: 1-5386-1786-2.

[100] Sergei Gerasenko et al. “Beacon Signals: What, Why, How, and Where?” In: *Computer* 34.10 (2001), pp. 108–110. ISSN: 0018-9162.

[101] *GISM - Global Ionospheric Scintillation Model*. 257. IEEA, Aug. 1, 2008.

[102] Andrea Goldsmith. *Wireless Communications*. Cambridge university press, 2005. ISBN: 0-521-83716-2.

[103] Gene H Golub and Charles F Van Loan. *Matrix Computations*. JHU press, 2013. ISBN: 1-4214-0859-7.

[104] Ian Goodfellow, Yoshua Bengio, and Aaron Courville. *Deep Learning*. Illustrated edição. Cambridge, Massachusetts: The MIT Press, Nov. 18, 2016. ISBN: 978-0-262-03561-3.

[105] Geoff Gordon. *10-725: Optimization.*

[106] GPS Silicon Valley. *GSV4004B User Manual*. Manual. Aug. 1, 2017.

[107] Ronald L Graham et al. "Concrete Mathematics: A Foundation for Computer Science". In: *Computers in Physics* 3.5 (1989), pp. 106–107. ISSN: 0894-1866.

[108] Simon Haykin and Barry Van Veen. *Signals and Systems*. Wiley, 1998. ISBN: 978-0-471-13820-4.

[109] Antony Hewish. "The Diffraction of Radio Waves in Passing through a Phase-Changing Ionosphere". In: *Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences* 209.1096 (1951), pp. 81–96.

[110] Are Hjorungnes and David Gesbert. "Complex-Valued Matrix Differentiation: Techniques and Key Results". In: *IEEE Transactions on Signal Processing* 55.6 (2007), pp. 2740–2746. ISSN: 1053-587X.

[111] Are Hjorungnes and David Gesbert. "Complex-Valued Matrix Differentiation: Techniques and Key Results". In: *IEEE Transactions on Signal Processing* 55.6 (2007), pp. 2740–2746. ISSN: 1053-587X.

[112] Bernhard Hofmann-Wellenhof, Herbert Lichtenegger, and James Collins. *Global Positioning System: Theory and Practice*. Springer Science & Business Media, 2012. URL: https://books.google.com/books?hl=en&lr=&id=F7jrCAAAQBAJ&oi=fnd&pg=PR19&dq=Global+Positioning+System:+Theory+and+Practice&ots=zYl4rXsOqD&sig=gFg3Pr5YHwnH8nxW6LcvRvuXSPw (visited on 03/12/2024).

[113] *Home · Convex.Jl*. URL: https://jump.dev/Convex.jl/stable/ (visited on 12/06/2022).

[114] Johannes Huber and Weilin Liu. "Data-Aided Synchronization of Coherent CPM-receivers". In: *IEEE Transactions on Communications* 40.1 (1992), pp. 178–189. ISSN: 0090-6778.

[115] Todd E Humphreys et al. "Data-Driven Testbed for Evaluating GPS Carrier Tracking Loops in Ionospheric Scintillation". In: *IEEE Transactions on Aerospace and Electronic Systems* 46.4 (2010), pp. 1609–1623. ISSN: 0018-9251.

[116] Todd E. Humphreys, Mark L. Psiaki, and Paul M. Kintner. "Modeling the Effects of Ionospheric Scintillation on GPS Carrier Phase Tracking". In: *IEEE Transactions on Aerospace and Electronic Systems* 46.4 (Oct. 2010), pp. 1624–1637. ISSN: 1557-9603. DOI: 10.1109/TAES.2010.5595583.

[117] Todd E. Humphreys et al. "Simulating Ionosphere-Induced Scintillation for Testing GPS Receiver Phase Tracking Loops". In: *IEEE Journal of Selected Topics in Signal Processing* 3.4 (Aug. 2009), pp. 707–715. ISSN: 1941-0484. DOI: 10.1109/JSTSP.2009.2024130.

[118] Vinay K Ingle and John G Proakis. *Digital Signal Processing Using MATLAB*. Cole Publishing Company, 2000.

[119] Hassan Ismail Fawaz et al. "Deep Learning for Time Series Classification: A Review". In: *Data mining and knowledge discovery* 33.4 (2019), pp. 917–963.

[120] Yu Jiao, John J Hall, and Yu T Morton. "Automatic Equatorial GPS Amplitude Scintillation Detection Using a Machine Learning Algorithm". In: *IEEE Transactions on Aerospace and Electronic Systems* 53.1 (2017), pp. 405–418. ISSN: 0018-9251.

[121] Yu Jiao, John Hall, and Yu Jade Morton. "Automatic GPS Phase Scintillation Detector Using a Machine Learning Algorithm". In: *Proceedings of the 2017 International Technical Meeting of The Institute of Navigation*. 2017, pp. 1160–1172.

[122] Yu Jiao et al. "Characteristics of Low-Latitude Signal Fading across the GPS Frequency Bands". In: Proceedings of the 27th International Technical Meeting of the Satellite Division of The Institute of Navigation (ION GNSS+ 2014). 2014, pp. 1203–1212. ISBN: 2331-5954.

[123] Yu Jiao et al. "Characterization of High-Latitude Ionospheric Scintillation of GPS Signals". In: *Radio Science* 48.6 (2013), pp. 698–708. ISSN: 1944-799X.

[124] Yu Jiao and Yu T Morton. "Comparison of the Effect of High-Latitude and Equatorial Ionospheric Scintillation on GPS Signals during the Maximum of Solar Cycle 24". In: *Radio Science* 50.9 (2015), pp. 886–903. ISSN: 1944-799X.

[125] Yu Jiao et al. "High Latitude Ionosphere Scintillation Characterization". In: Institute of Navigation International Technical Meeting 2013, ITM 2013. Jan. 1, 2013, pp. 579–584.

[126] Yu Jiao. "Low-Latitude Ionospheric Scintillation Signal Simulation, Characterization, and Detection on GPS Signals". thesis. Colorado State University, 2017.

[127] Yu Jiao, John J. Hall, and Yu T. Morton. "Performance Evaluation of an Automatic GPS Ionospheric Phase Scintillation Detector Using a Machine-Learning Algorithm". In: *NAVIGATION* 64.3 (2017), pp. 391–402. ISSN: 2161-4296. DOI: 10.1002/navi.188. URL: https://onlinelibrary.wiley.com/doi/abs/10.1002/navi.188 (visited on 05/02/2023).

[128] Elliott D. Kaplan and Christopher Hegarty. *Understanding GPS/GNSS: Principles and Applications*. Artech house, 2017.

[129] Steven M. Kay. *Fundamentals of Statistical Processing, Volume 2: Detection Theory*. Pearson Education India, 2009.

[130] Steven M. Kay. *Fundamentals of Statistical Signal Processing: Estimation Theory*. Prentice-Hall, Inc., 1993.

[131] MC Kelley. "Equatorial Spread-F: Recent Results and Outstanding Problems". In: *Journal of atmospheric and terrestrial physics* 47.8-10 (1985), pp. 745–752. ISSN: 0021-9169.

[132] P. M. Kintner et al. "Size, Shape, Orientation, Speed, and Duration of GPS Equatorial Anomaly Scintillations". In: *Radio Science* 39.2 (2004). ISSN: 1944-799X. DOI: 10.1029/2003RS002878. URL: https://onlinelibrary.wiley.com/doi/abs/10.1029/2003RS002878 (visited on 09/06/2023).

[133] Mark Knight and Anthony Finn. "The Effects of Ionospheric Scintillations on GPS". In: *Proceedings of the 11th International Technical Meeting of the Satellite Division of the Institute of Navigation (ION GPS 1998)*. 1998, pp. 673–685.

[134] Mark F Knight, Anthony Finn, and Manuel Cervera. *Ionospheric Effects on Global Positioning System Receivers*. DEFENCE SCIENCE and TECHNOLOGY ORGANISATION CANBERRA (AUSTRALIA), 1998.

[135] Mark Frederick Knight. "Ionospheric Scintillation Effects on Global Positioning System Receivers". thesis. University of Adelaide, 2000.

[136] Tamara G. Kolda and Brett W. Bader. "Tensor Decompositions and Applications". In: *SIAM review* 51.3 (2009), pp. 455–500.

[137] Basil Kouvaritakis and Mark Cannon. "Model Predictive Control". In: *Switzerland: Springer International Publishing* 38 (2016).

[138] Erwin Kreyszig, K Stroud, and G Stephenson. *Advanced Engineering Mathematics*. Vol. 9. John Wiley & Sons, Inc. 9 th edition, 2006 Page 2 of 6 Teaching methods ..., 2008.

[139] P.A. Kullstam and M.J. Keskinen. "Ionospheric Scintillation Effects on UHF Satellite Communications". In: *MILCOM 2000 Proceedings. 21st Century Military Communications. Architectures and Technologies for Information Superiority (Cat. No.00CH37155)*. MILCOM 2000 Proceedings. 21st Century Military Communications. Architectures and Technologies for Information Superiority (Cat. No.00CH37155). Vol. 2. Oct. 2000, 779–783 vol.2. DOI: 10.1109/MILCOM.2000.904036.

[140] B. P. Lathi and Adel S. Sedra. *Linear Systems and Signals*. New York Oxford, July 1, 2004. 973 pp. ISBN: 978-0-19-515833-5.

[141] Bhagwandas P Lathi. *Modern Digital and Analog Communication Systems*. Oxford University Press, Inc., 1995. ISBN: 0-03-028407-4.

[142] Alberto Leon-Garcia. *Probability, Statistics, and Random Processes for Electrical Engineering*. 3rd ed. edição. Upper Saddle River, NJ: Prentice Hall, 2007. ISBN: 978-0-13-147122-1.

[143] Nicola Linty et al. "Detection of GNSS Ionospheric Scintillations Based on Machine Learning Decision Tree". In: *IEEE Transactions on Aerospace and Electronic Systems* 55.1 (2018), pp. 303–317. ISSN: 0018-9251.

[144] Yunxiang Liu, Y Jade Morton, and Yu Jiao. "Application of Machine Learning to the Characterization of GPS L1 Ionospheric Amplitude Scintillation". In: 2018 IEEE/ION Position, Location and Navigation Symposium (PLANS). IEEE, 2018, pp. 1159–1166. ISBN: 1-5386-1647-5.

[145] Sergi Locubiche-Serra, Gonzalo Seco-Granados, and José A. López-Salcedo. "Doubly-Adaptive Autoregressive Kalman Filter for GNSS Carrier Tracking under Scintillation Conditions". In: June 1, 2016, pp. 1–6. DOI: 10.1109/ICL-GNSS.2016.7533859.

[146] Rafael Lopes et al. "Linear Time-Invariant Filtering for Real-Time Monitoring of Ionospheric Scintillation in GNSS Receivers". In: *GPS Solutions* 27 (May 20, 2023). DOI: 10.1007/s10291-023-01470-0.

[147] Antônio MP Lucena et al. "Flexible FPGA-based BPSK Signal Generator for Space Applications". In: *International Journal of Circuits, Systems and Signal Processing* 8 (2014), pp. 160–165.

[148] Christophe Macabiau et al. "Kalman Filter Based Robust GNSS Signal Tracking Algorithm in Presence of Ionospheric Scintillations". In: *Proceedings of the 25th International Technical Meeting of The Satellite Division of the Institute of Navigation (ION GNSS 2012)*. 2012, pp. 3420–3434.

[149] J. W MacDougall. "Distributions of the Irregularities Which Produce Ionospheric Scintillations". In: *Journal of Atmospheric and Terrestrial Physics* 43.4 (Apr. 1, 1981), pp. 317–325. ISSN: 0021-9169. DOI: 10.1016/0021-9169(81)90093-3. URL: https://www.sciencedirect.com/science/article/pii/0021916981900933 (visited on 08/18/2023).

[150] Tarcisio F Maciel. "Slides - Otimização não-linear".

[151] Mamatha R Maheshwarappa and Christopher P Bridges. "Software Defined Radios for Small Satellites". In: 2014 NASA/ESA Conference on Adaptive Hardware and Systems (AHS). IEEE, 2014, pp. 172–179. ISBN: 1-4799-5356-3.

[152] Sylvie Marcos et al. "A Unified Framework for Gradient Algorithms Used for Filter Adaptation and Neural Network Training". In: *International Journal of Circuit Theory and Applications* 20.2 (1992), pp. 159–200. ISSN: 1097-007X. DOI: 10.1002/cta.4490200205. URL: https://onlinelibrary.wiley.com/doi/abs/10.1002/cta.4490200205 (visited on 05/29/2023).

[153] Umberto Mengali. "Joint Phase and Timing Acquisition in Data-Transmission". In: *IEEE Transactions on Communications* 25.10 (1977), pp. 1174–1185. ISSN: 0090-6778.

[154] Umberto Mengali. *Synchronization Techniques for Digital Receivers*. Springer Science & Business Media, 1997.

[155] Heinrich Meyr, Marc Moeneclaey, and Stefan A. Fechtel. *Digital Communication Receivers: Synchronization, Channel Estimation, and Signal Processing*. Vol. 444. Wiley Online Library, 1998.

[156] Thomas P. Minka. "Old and New Matrix Algebra Useful for Statistics". In: *See www. stat. cmu. edu/minka/papers/matrix. html* 4 (2000). URL: https://citeseerx.ist.psu.edu/document?repid=rep1&type=pdf&doi=82d147ee42075cc1d0b2509d2767d14329b193e5 (visited on 03/16/2024).

[157] Sridhar Miriyala, Padma Raju Koppireddi, and Srinivasa Rao Chanamallu. "Robust Detection of Ionospheric Scintillations Using MF-DFA Technique". In: *Earth, Planets and Space* 67.1 (2015), pp. 1–5. ISSN: 1880-5981.

[158] Pratap Misra and Per Enge. "Global Positioning System: Signals, Measurements and Performance (Lincoln, MA: Ganga". In: *Global Positioning System: Signals, Measurements and Performance Lincoln, MA: Ganga* (2006).

[159] Todd K Moon and Wynn C Stirling. *Mathematical Methods and Algorithms for Signal Processing*. 2000. ISBN: 0-201-36186-8.

[160] Michele Morelli, Umberto Mengali, and Giorgio Matteo Vitetta. "Joint Phase and Timing Recovery with CPM Signals". In: *IEEE Transactions on Communications* 45.7 (1997), pp. 867–876. ISSN: 0090-6778.

[161] Minoru Nakagami. "The M-Distribution—A General Formula of Intensity Distribution of Rapid Fading". In: *Statistical Methods in Radio Wave Propagation*. Elsevier, 1960, pp. 3–36.

[162] L. J. Nickisch. "Practical Applications of Haselgrove's Equations for HF Systems". In: *URSI Radio Science Bulletin* 2008.325 (June 2008), pp. 36–48. ISSN: 1024-4530. DOI: 10.23919/URSIRSB.2008.7909584.

[163] Josef Nossek. *Adaptive and Array Signal Processing*. 2015.

[164] "A Note on Scintillation Indices". In: *Planetary and Space Science* 17.5 (May 1, 1969), pp. 1059–1066. ISSN: 0032-0633. DOI: 10.1016/0032-0633(69)90112-3. URL: https://www.sciencedirect.com/science/article/pii/0032063369901123 (visited on 04/28/2023).

[165] Fernando D Nunes and Fernando MG Sousa. "Generation of Nakagami Correlated Fading in GNSS Signals Affected by Ionospheric Scintillation". In: 2014 7th ESA Workshop on Satellite Navigation Technologies and European Workshop on GNSS Signals and Signal Processing (NAVITEC). IEEE, 2014, pp. 1–8. ISBN: 1-4799-6529-4.

[166] Fernando D Nunes and Fernando MG Sousa. "Practical Simulation of GNSS Signals in the Presence of Ionospheric Scintillation". In: 2014 IEEE/ION Position, Location and Navigation Symposium-PLANS 2014. IEEE, 2014, pp. 50–58. ISBN: 1-4799-3320-1.

[167] Alan V. Oppenheim and Ronald W. Schafer. *Discrete-Time Signal Processing: International Edition*. 3ª edição. Upper Saddle River Munich: Pearson, Nov. 12, 2009. ISBN: 978-0-13-206709-6.

[168] Alan Oppenheim, Alan Willsky, and S. Nawab. *Signals and Systems*. 2º edição. Upper Saddle River, N.J: Pearson, 1996. ISBN: 978-0-13-814757-0.

[169] Athanasios Papoulis et al. *Probability, Random Variables and Stochastic Processes*. 4th ed. edição. Boston: McGraw-Hill Education, 2001. ISBN: 978-0-07-366011-0.

[170] Kaare Brandt Petersen and Michael Syskind Pedersen. "The Matrix Cookbook". In: *Technical University of Denmark* 7.15 (2008), p. 510.

[171] Shishir Priyadarshi. "A Review of Ionospheric Scintillation Models". In: *Surveys in geophysics* 36 (2015), pp. 295–324.

[172] John G. Proakis. *Digital Communications*. 4th ed. Mc Graw Hill, Jan. 1, 2000. 1024 pp. ISBN: 978-0-07-118183-9. URL: https://www.amazon.com/Digital-Communications-Proakis/dp/0071181830.

[173] John Proakis and Masoud Salehi. *Digital Communications*. 5th ed. Boston: Mc Graw Hill, Jan. 1, 2007. ISBN: 978-0-07-295716-7.

[174] Mark L. Psiaki and Hee Jung. "Extended Kalman Filter Methods for Tracking Weak GPS Signals". In: *Proceedings of the 15th International Technical Meeting of the Satellite Division of the Institute of Navigation (ION GPS 2002)*. 2002, pp. 2539–2553.

[175] Simon Ramo, John R. Whinnery, and Theodore Van Duzer. *Fields and Waves in Communication Electronics*. John Wiley & Sons, 1994.

[176] Simon Ramo, John R Whinnery, and Theodore Van Duzer. *Fields and Waves in Communication Electronics*. John Wiley & Sons, 1994. ISBN: 81-265-1525-2.

[177] John Ashworth Ratcliffe. "Some Aspects of Diffraction Theory and Their Application to the Ionosphere". In: *Reports on progress in physics* 19.1 (1956), p. 188. ISSN: 0034-4885.

[178] *RECOMMENDATION ITU-R P.838-3 - Specific Attenuation Model for Rain for Use in Prediction Methods*.

[179] P.A. Regalia. "Fundamentals of Adaptive Filtering [Book Review]". In: *IEEE Control Systems Magazine* 25.4 (Aug. 2005), pp. 77–79. ISSN: 1941-000X. DOI: 10.1109/MCS.2005.1499393.

[180] J. M. Retterer. "Forecasting Low-Latitude Radio Scintillation with 3-D Ionospheric Plume Models: 2. Scintillation Calculation". In: *Journal of Geophysical Research: Space Physics* 115.A3 (2010).

[181] CL Rino and EJ Fremouw. "The Angle Dependence of Singly Scattered Wavefields". In: *Journal of Atmospheric and Terrestrial Physics* 39.8 (1977), pp. 859–868. ISSN: 0021-9169.

[182] Charles L. Rino et al. "A Characterization of Intermediate-Scale Spread F Structure from Four Years of High-Resolution C/NOFS Satellite Data". In: *Radio Science* 51.6 (2016), pp. 779–788.

[183] CHARLES L. Rino and CHARLES S. Carrano. “A Compact Strong-Scatter Scintillation Model”. In: *Proceedings of the International Beacon Satellite Symposium*. Citeseer, 2013.

[184] C. L. Rino. “Some New Results on the Statistics of Radio Wave Scintillation, 2. Scattering from a Random Ensemble of Locally Homogeneous Patches”. In: *Journal of Geophysical Research* 81.13 (1976), pp. 2059–2064.

[185] C. L. Rino, R. C. Livingston, and H. E. Whitney. “Some New Results on the Statistics of Radio Wave Scintillation, 1. Empirical Evidence for Gaussian Statistics”. In: *Journal of Geophysical Research* 81.13 (1976), pp. 2051–2057.

[186] C. L. Rino. “A Power Law Phase Screen Model for Ionospheric Scintillation: 2. Strong Scatter”. In: *Radio Science* 14.6 (1979), pp. 1147–1155.

[187] CL Rino. “A Power Law Phase Screen Model for Ionospheric Scintillation: 1. Weak Scatter”. In: *Radio science* 14.6 (1979), pp. 1135–1145. ISSN: 1944-799X.

[188] C. L. Rino, C. S. Carrano, and Patrick Roddy. “Wavelet-Based Analysis and Power Law Classification of C/NOFS High-Resolution Electron Density Data”. In: *Radio Science* 49.8 (2014), pp. 680–688.

[189] Kenneth H Rosen. “Discrete Mathematics and Its Applications (7Th Editio)”. In: *William C Brown Pub* (2011).

[190] Clifford L. Rufenach. “Wavelength Dependence of Radio Scintillation: Ionosphere and Interplanetary Irregularities”. In: *Journal of Geophysical Research* 79.10 (1974), pp. 1562–1566.

[191] *Rules*. URL: https://dcp.stanford.edu/rules (visited on 12/08/2022).

[192] Salman Sadruddin and M Sohail. “FPGA Based Tele Command ReceiverModule for Microsatellites”. In: *open journal of communications and software* (2014).

[193] Peter J Schreier and Louis L Scharf. *Statistical Signal Processing of Complex-Valued Data: The Theory of Improper and Noncircular Signals*. Cambridge university press, 2010. ISBN: 1-139-48762-0.

[194] Shayle R Searle and Andre I Khuri. *Matrix Algebra Useful for Statistics*. John Wiley & Sons, 2017. ISBN: 1-118-93514-4.

[195] J. A. Secan et al. “High-Latitude Upgrade to the Wideband Ionospheric Scintillation Model”. In: *Radio Science* 32.4 (1997), pp. 1567–1574.

[196] J. A. Secan et al. “An Improved Model of Equatorial Scintillation”. In: *Radio science* 30.3 (1995), pp. 607–617.

[197] Septentrio. *PolaRx5S User Manual*. 1.2. Manual. Mar. 15, 2018.

[198] P. Shaft. “On the Relationship Between Scintillation Index and Rician Fading”. In: *IEEE Transactions on Communications* 22.5 (May 1974), pp. 731–732. ISSN: 1558-0857. DOI: 10.1109/TCOM.1974.1092244.

[199] V. I. Shishov. “Dependence of the Form of the Scintillation Spectrum on the Form of the Spectrum of Refractive-Index Inhomogeneities Communication I. A Phase Screen”. In: *Radiophysics and Quantum Electronics* 17.11 (1974), pp. 1287–1292.

[200] I. P. Shkarofsky. “Generalized Turbulence Space-Correlation and Wave-Number Spectrum-Function Pairs”. In: *Canadian Journal of Physics* 46.19 (1968), pp. 2133–2153.

[201] James J. Spilker Jr et al. *Global Positioning System: Theory and Applications, Volume I*. American Institute of Aeronautics and Astronautics, 1996.

[202] James Stewart. *Calculus*. Cengage Learning, 2011. ISBN: 1-133-17069-2.

[203] Gilbert Strang et al. *Introduction to Linear Algebra*. Vol. 3. Wellesley-Cambridge Press Wellesley, MA, 1993.

[204] Gilbert Strang. *Linear Algebra and Learning from Data*. Vol. 4. Wellesley-Cambridge Press Cambridge, 2019.

[205] J Sanz Subirana, J M Juan Zornoza, and M Hernández-Pajares. *GNSS Data Processing, Vol. I: Fundamentals and Algorithms*. Vol. 1. ESA Communications. ISBN: 978-92-9221-886-7.

[206] J Sanz Subirana, J M Juan Zornoza, and M Hernández-Pajares. *GNSS Data Processing, Vol. II: Laboratory Exercises*. Vol. 2. ESA Communications. ISBN: 978-92-9221-886-7.

[207] PJ Sultan. "Linear Theory and Modeling of the Rayleigh-Taylor Instability Leading to the Occurrence of Equatorial Spread F". In: *Journal of Geophysical Research: Space Physics* 101.A12 (1996), pp. 26875–26891. ISSN: 0148-0227.

[208] Valerian Ilich Tatarski. *Wave Propagation in a Turbulent Medium*. Courier Dover Publications, 2016.

[209] Peter JG Teunissen and Oliver Montenbruck. *Springer Handbook of Global Navigation Satellite Systems*. Vol. 10. Springer, 2017.

[210] Sergios Theodoridis. *Machine Learning: A Bayesian and Optimization Perspective*. 2nd ed. Academic Pr, 2020. ISBN: 978-0-12-818803-3.

[211] Sarang Thombre and Jari Nurmi. "Bandpass-Sampling Based GNSS Sampled Data Generator - A Design Perspective". In: *2012 International Conference on Localization and GNSS*. 2012 International Conference on Localization and GNSS (ICL-GNSS). Starnberg, Germany: IEEE, June 2012, pp. 1–6. ISBN: 978-1-4673-2343-7 978-1-4673-2344-4 978-1-4673-2342-0. DOI: 10.1109/ICL-GNSS.2012.6253114. URL: http://ieeexplore.ieee.org/document/6253114/ (visited on 01/19/2023).

[212] *Transionospheric Radio Propagation. The Global Ionospheric Scintillation Model (GISM)*. Rep. ITU-R P.2097. 2007.

[213] Roland T. Tsunoda. "High-Latitude F Region Irregularities: A Review and Synthesis". In: *Reviews of Geophysics* 26.4 (1988), pp. 719–760.

[214] Frank Vahid. *Digital Design with RTL Design, VHDL, and Verilog*. John Wiley & Sons, Mar. 9, 2010. 592 pp. ISBN: 978-0-470-53108-2.

[215] AJ Van Dierendonck, John Klobuchar, and Quyen Hua. "Ionospheric Scintillation Monitoring Using Commercial Single Frequency C/A Code Receivers". In: Proceedings of ION GPS. Vol. 93. 1993, pp. 1333–1342.

[216] A. J. Van Dierendonck and B. Arbesser-Rastburg. "Measuring Ionospheric Scintillation in the Equatorial Region Over Africa, Including Measurements From SBAS Geostationary Satellite Signals". In: Proceedings of the 17th International Technical Meeting of the Satellite Division of The Institute of Navigation (ION GNSS 2004). Sept. 24, 2004, pp. 316–324. URL: http://www.ion.org/publications/abstract.cfm?jp=p&articleID=5710 (visited on 04/16/2023).

[217] Harry L. Van Trees. *Detection, Estimation, and Modulation Theory, Part I: Detection, Estimation, and Linear Modulation Theory*. John Wiley & Sons, 2004.

[218] Harry L Van Trees. *Optimum Array Processing: Part IV of Detection, Estimation, and Modulation Theory*. John Wiley & Sons, 2002. ISBN: 0-471-09390-4.

[219] K. Vasudevan, K. Giridhar, and Bhaskar Ramamurthi. "Nyquist-Rate Detection of CPM Signals Using the Viterbi Algorithm". In: *Proc. of the Third National Conf. on Commun., Indian Institute of Technology Madras, India*. 1997, pp. 85–90. URL: http://home.iitk.ac.in/~vasu/nc970.pdf (visited on 03/08/2024).

[220] Dmytro Vasylyev et al. *Global Ionospheric Scintillation Model: Current Status and Further Development Strategies*. other. pico, Mar. 4, 2021. DOI: 10.5194/egusphere-egu21-9441. URL: https://meetingorganizer.copernicus.org/EGU21/EGU21-9441.html (visited on 08/14/2023).

[221] Jordi Vilà-Valls, José López-Salcedo, and Gonzalo Seco-Granados. *An Interactive Multiple Model Approach for Robust GNSS Carrier Phase Tracking under Scintillation Conditions — IEEE Conference Publication — IEEE Xplore*. URL: https://ieeexplore.ieee.org/document/6638896 (visited on 02/13/2024).

[222] Jordi Vilà-Valls et al. "Survey on Signal Processing for GNSS under Ionospheric Scintillation: Detection, Monitoring, and Mitigation". In: *NAVIGATION: Journal of the Institute of Navigation* 67.3 (2020), pp. 511–535. ISSN: 0028-1522.

[223] HE Whitney et al. "Estimation of the Cumulative Amplitude Probability Distribution Function of Ionospheric Scintillations". In: *Radio Science* 7.12 (1972), pp. 1095–1104. ISSN: 1944-799X.

[224] H. E. Whitney, J. Aarons, and C. Malik. “A Proposed Index for Measuring Ionospheric Scintillations”. In: *Planetary and Space Science* 17.5 (May 1, 1969), pp. 1069–1073. ISSN: 0032-0633. DOI: 10.1016/0032-0633(69)90114-7. URL: https://www.sciencedirect.com/science/article/pii/0032063369901147 (visited on 04/26/2023).

[225] Zhong Ye and H. Satorius. “Channel Modeling and Simulation for Mobile User Objective System (MUOS)–Part 1: Flat Scintillation and Fading”. In: *IEEE International Conference on Communications, 2003. ICC '03.* IEEE International Conference on Communications, 2003. ICC '03. Vol. 5. May 2003, 3503–3510 vol.5. DOI: 10.1109/ICC.2003.1204106.

[226] Kung Chie Yeh and Chao-Han Liu. “Radio Wave Scintillations in the Ionosphere”. In: *Proceedings of the IEEE* 70.4 (1982), pp. 324–360. ISSN: 0018-9219.

[227] QT Zhang. “A Decomposition Technique for Efficient Generation of Correlated Nakagami Fading Channels”. In: *IEEE Journal on Selected Areas in Communications* 18.11 (2000), pp. 2385–2392. ISSN: 0733-8716.

[228] Rodger Ziemer and William H Tranter. *Principles of Communications: System Modulation and Noise*. John Wiley & Sons, 2006. ISBN: 81-265-0839-6.