A Novel Class of Intensity-based Metrics for Image Functions which Accommodate a Generalized Weber's Model of Perception

by

Dongchang Li

A thesis
presented to the University of Waterloo
in fulfillment of the
thesis requirement for the degree of
Master of Mathematics
in
Applied Mathematics

Waterloo, Ontario, Canada, 2020

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Abstract

Even though the usual L^2 image fidelity measurements, such as MSE and PSNR, characterize the mean error for each pixel of the images, these traditional measurements are not designed to predict human visual perception of image quality. In other words, the standard L^2-type optimization in the context of best approximation theory is not in accordance with human visual system. To provide alternative methods of measuring image distortion perceptually, the structural similarity image quality measure (SSIM) was established decades ago.

In this thesis, we are concerned with constructing a novel class of metrics via intensity-based measures, which accommodate a well-known psychological model, Weber's model of perception, by allowing greater deviations at higher intensity values and lower deviations at lower intensity values. The standard Weber model, however, is known to fail at low and high intensities, which has given rise to a generalized class of Weber models. We have derived a set of "Weberized" distance functions which accommodate these generalized models. Mathematically, we prove the existence and uniqueness of the density functions associated with the measures which conform to generalized Weber's model of perception. Meanwhile, we consider the generalized Weber-based metrics as the optimization criteria and implement them in best approximation problems, where we also prove the existence and uniqueness of best approximations. We compare the results, which are theoretically adapted to the human visual system, with best L^2-based approximations. Using a functional analysis point of view, we examine the stationary equations associated with the generalized Weber-based metrics to arrive at the Fréchet derivatives of these metrics. Finally, we establish an existence-uniqueness theorem of best generalized Weber-based L^2 approximations in finite-dimensional Hilbert spaces.

Acknowledgements

First of all, I would like to thank my supervisor, Prof. Edward R. Vrscay, for his dedication, guidance, support and patience throughout both of my undergraduate and master studies. Indeed, Prof. Vrscay is really the one who introduced me to the research in the field of mathematical imaging, and taught me academic skills in terms of writing, presentation and so on. Over the last a few years of my study, he has not only kindly suggested me to take diverse graduate courses in both aspects of mathematics and engineering, but raised insightful and profound ideas which helped me with my research projects. More importantly, the way in which he conducts research and takes care of his students has significant impacts on my future career and life.

Additionally, I wish to thank my co-supervisor and co-author, Prof. Davide La Torre, for providing invaluable suggestions and comments on my master's research, and your feedback about my technical questions has been extremely helpful. My thanks are also sent to Prof. Giang Tran and Prof. Henry Wolkowicz for their willingness to be the readers of my thesis and the committee members of my Master's defence. To Giang, thank you for taking the time to talk to me when I have been considering my future study, and I sincerely appreciate your encouragement and advice. I would also like to express my gratitude to Henry for providing much help on my best approximation problems, especially when I first started learning convex optimization.

Finally, the appreciation goes to my office mates, Ming Miao and Aiden Huffman, and my colleagues around the Department of Applied Mathematics at the University of Waterloo, in particular, Zhibing Sun and William Bell. Thank you for all the support, encouragement and joy you gave to me while I was writing this thesis.

Dedication

This thesis is dedicated to my family.

Table of Contents

List of Figures

Chapter 1

Introduction

In the context of mathematical imaging, one of the most interesting problems is to construct the metrics or distance functions which can appropriately measure the similarity or difference between two images. In practice, these metrics may be applied into many fundamental image processing problems, including, but not limited to, image quality assessment [36, 37], image compression [35] and image reconstruction [39]. While there are some efforts that demonstrate how to train a visibility metric using deep learning [38], mathematically, there are few distance functions which are capable of describing how "close" the two images (or signals) are in terms of the human visual system (HVS). Fortunately, there is a traditional law, so-called Weber's Law, which accounts for the relation between the relative change of stimulus and the background stimulus. In the application of our purpose, this stimulus refers to the human vision sensation, which motivated us to establish the Weber's model of visual perception [7]. In [1], the authors presented one method of modifying the usual L^2 distance metric so that they are in accordance with Weber's model of perception. Moreover, it has been proved that the most natural "Weberized" image metric is the logarithmic L^1 distance, from which one may obtain the generalization of "Weberized" metrics, i.e., logarithmic L^p distances [1].

The motivation of this thesis is to build up a novel class of metrics via intensity-based measures, which accommodate Weber's model of perception—allowing greater deviations at higher intensity values and lower deviations at lower intensity values. Indeed, we have generalized the Weber's model of perception and derived the generalized form of "Weberized" distance functions. In the meantime, we are extremely interested in the best approximation problems in the finite dimensional Hilbert space, in which the optimization is done with respect to the generalized Weber-based L^2 distance functions and the structural similarity (SSIM) function. In the greyscale range space, we prove the existence and

uniqueness of the density functions associated with the intensity-based measures, which conform to generalized Weber's model of perception, by a powerful functional equation: Abel's equation [40].

The remainder of this thesis is structured as follows. Chapter 2 contains the mathematical background required to understand the representation of images as functions and implement the best image functions approximation. Weber's model of perception, coming from Weber's law, is firstly presented. Then, we review some fundamental basics of signal and image processing, including the discrete cosine transform (DCT), which was used to test on experiments of image function approximation using different types of distance metrics. Furthermore, we provide an elementary acquaintance for two typical image fidelity measurements—mean squared error (MSE) and structural similarity (SSIM).

In Chapter 3, we first review previous efforts to Weberize the distance functions which involve the ratio between two image functions. We then show how distance functions can be Weberized, in terms of a generalized Weber's model of perception, by employing nonuniform measures on greyscale range space. In addition, the leading order terms of generalized Weber-based distance functions are exploited in this chapter.

As for Chapter 4, we focus on the mathematical aspects of our problem, i.e., existence, uniqueness and asymptotic behaviours of the density functions associated with intensity-based measures conforming to generalized Weber's model of perception. Essentially, the Abel's equation and iteration dynamics of the density functions play a key role in proving the above mathematical properties of the density functions.

In Chapter 5, we present a rigorous analysis of the stationarity conditions on the generalized Weber-based and SSIM-based approximations. In addition, we investigate the Frećhet derivatives of the distance functionals utilized in optimization algorithms in the two aforementioned approximation methods.

Chapter 6 consists of the existence and uniqueness of the optimal solution for the best approximation of non-negative functions via a standard L^2 metric and a set of Weber-based L^2 metrics. In practice, we implement the Gauss-Newton algorithm to obtain a best affine logarithmic L^2 approximation in finite dimensional Hilbert space.

Finally, we summarize the contributions of this thesis and provide some possible extensions of this work in Chapter 7.

Chapter 2

Background

This chapter introduces some background information to help the readers to understand mathematical preliminaries of this thesis. In what follows, we state Weber's Law and then describe Weber's model of perception and its generalization. Furthermore, we show how to represent the signals and images as 1-dimensional and 2-dimensional functions in terms of both continuous and discrete versions. Last, but not least, a number of traditional and well-known concepts related to signal and image processing, including discrete cosine transform (DCT), mean squared error (MSE) and structural similarity (SSIM), have been provided in this chapter.

2.1 Weber's Model of Perception and its Generalization

Weber's law, a well-known perceptual law in the field of psychology and psychophysics, was originally described by German physiologist Ernst Heinrich Weber [18], and was then developed by his student Gustav Theodor Fechner [19]. According to Weber's law, the change of a human being's sensitivity can be quantified by the background physical stimulus. More specifically, the smallest change in the stimulus that produces a just noticeable difference, i.e., ΔI, is proportional to the initial stimulus I, i.e.,

$$\frac{\Delta I}{I} = k, \tag{2.1}$$

where k, so-called Weber's fraction, is a constant. Note that the stimulus can be in any of the human sensations, e.g., vision, hearing, etc.

In reality, all images such as greyscale and colour images are perceived by human vision. With reference to the human visual system, we employ the effect of Weber's law to establish Weber's model of visual perception: Consider a image function $u(x)$ which represents the intensity at a point x. Then the minimum perceived sensitivity Δu is related to the intensity u as follows,

$$\frac{\Delta u}{u} = c, \tag{2.2}$$

where c is a constant over a significant range of u-values. In practice, this standard Weber's law in Eq. (2.2) has been applied in some image and visual processing methods, such as total variation (TV) restoration [7], Mumford-Shah segmentation [8] and visualization discrimination [9]. The main idea of these methods is to "Weberize" – modify the image or visual processors so that they accommodate Weber's model of perception. However, in some situations, e.g., luminance and contrast discrimination [10, 11] and duration discrimination [12], a generalized form of Weber's law is employed to construct the perception model. This motivates us to investigate on a generalized Weber's model of perception: Given a greyscale background intensity $I > 0$, the minimum change in intensity ΔI perceived by the human visual system (HVS) is related to I as follows,

$$\frac{\Delta I}{I^a} = C, \tag{2.3}$$

where $a > 0$ and C is constant, or at least roughly constant, over a significant range of intensity values I. Note that the special case $a = 1$ corresponds to Weber's standard model, and the case $a = 0$ corresponds to an absence of Weber's model. In essence, Eqs. (2.2) and (2.3) show that the human visual system (HVS) is more (less) sensitive to changes at which greyscale intensities (i.e., $|u(x)|$) are small (large) [1]. In other words, a generalized Weber-based distance between two functions u and v should allow lesser deviations at lower intensity values and greater deviations at higher intensity values.

2.2 Representation of Images as Functions

Realistically, a 2-dimensional (2D) image can be thought of as a function of two real spatial variables x and y, say $f(x, y)$, where x and y are continuous real variables defining points of the image, and $f(x, y)$, typically a real value, represents the intensity of the image at point (x, y). Mathematically, $f(x, y)$, i.e., the image function, is described as follows,

$$f : \mathbb{R}^2 \to \mathbb{R}. \tag{2.4}$$

As such, a number of tasks for analyzing images can be achieved by appropriate functional analysis. For greyscale images, we may define the domain of $f(x, y)$ as $X \subset \mathbb{R}^2$, which is usually called the pixel space. Without loss of generality, we can assume that $X = [a, b]^2 \subset \mathbb{R}^2$, and the range of f is the interval $[A, B]$, where a, b, A and B are non-negative constants. Then, the value A will represent black and B will represent white. An intermediate value, i.e., $A < f(x, y) < B$, will represent some shaded grey value. In many studies, the greyscale space $[A, B]$ is normalized to the interval $[0, 1]$.

Indeed, Eq. (2.4) states the continuous version of image functions. However, one most often works with discrete image functions in practice, which are essentially interpreted as digital images. A digital image is a 2D image $g(x, y)$ described by a discrete 2D array of intensity samples. Given any greyscale digital image, it is represented by an m by n matrix, denoted as $\mathbf{u} = \{u_{ij}\}$, where i and j are positive integers, and $1 \leq i \leq m$, $1 \leq j \leq n$. Moreover, each entry u_{ij} of the matrix represents the greyscale values at the (i, j) pixel. The intensity range of the digital image is also quantized into a discrete values, which allow the image to be easily stored in digital memory. As a popular case of 8-bit images, each greyscale value is recorded as a number in the set $\{0, 1, 2, \cdots, 255\}$.

2.3 Discrete Cosine Transform (DCT)

The discrete cosine transform (DCT) was originally invented by Nasir Ahmed in 1972 for image compression [20]. Since then, it has been further developed and widely applied in many image and video compression standards such as JPEG and MPEG. In fact, the main idea of discrete cosine transform is to describe the finite set of discrete data as a weighted sum of cosine functions. Similar to discrete Fourier transform (DFT), DCT converts a signal or image from the spatial domain to frequency domain. There actually exist a few different versions of DCT, and the most commonly used form is so-called "DCT-II" method, introducing the periodic extension of $2N$-point data set $f[n]$ which, in turn, is an even extension of the N-point data set $(f[0], f[1], \cdots, f[N-1])$ by reflecting with respect to $n = -1/2$,

$$F[k] = \sum_{n=0}^{N-1} f[n] \cos\left(\frac{\pi}{N}\left(n + \frac{1}{2}\right)k\right), \quad k = 0, 1, \cdots, N-1. \tag{2.5}$$

Now, if we let the set of N-vectors u_k, $k = 0, 1, \cdots, N-1$, with components defined as follows,

$$\phi_k[n] = \cos\left(\frac{\pi}{N}\left(n + \frac{1}{2}\right)k\right), \quad n = 0, 1, \cdots, N-1, \tag{2.6}$$

it is straightforward to obtain that,

$$\langle \phi_0, \phi_0 \rangle = N, \quad \langle \phi_k, \phi_k \rangle = \frac{N}{2}, \quad k \neq 0, \tag{2.7}$$

and

$$\langle \phi_k, \phi_l \rangle = 0, \quad k \neq l. \tag{2.8}$$

In other words, the N-vectors ϕ_k form an orthogonal set in $\mathbb{R}^N$. From Eqs. (2.7) and (2.8), by normalization process, it follows that the set of N-vectors, e_k, defined below, forms an orthonormal basis on $\mathbb{R}^N$:

$$e_k[n] = \lambda_k \sqrt{\frac{2}{N}} \cos\left(\frac{\pi}{N}\left(n + \frac{1}{2}\right)k\right), \quad n = 0, 1, \cdots, N-1, \tag{2.9}$$

where

$$\lambda_k = \begin{cases} \frac{1}{\sqrt{2}}, & k = 0 \\ 1, & k \neq 0. \end{cases} \tag{2.10}$$

Given any $f \in \mathbb{R}^N$, we can express it as the linear expansion in terms of the orthonormal basis e_k, i.e.,

$$f = \sum_{n=0}^{N-1} c_k e_k, \tag{2.11}$$

where c_k, usually called the Fourier coefficients of f, are given by

$$c_k = \langle f, e_k \rangle = \sum_{n=0}^{N-1} f[n] \lambda_k \sqrt{\frac{2}{N}} \cos\left(\frac{\pi}{N}\left(n + \frac{1}{2}\right)k\right). \tag{2.12}$$

With the understanding of Fourier transform, we consider the c_k to define the discrete cosine transform (DCT) of f, i.e.,

$$F[k] = \lambda_k \sqrt{\frac{2}{N}} \sum_{n=0}^{N-1} f[n] \cos\left(\frac{\pi}{N}\left(n + \frac{1}{2}\right)k\right). \tag{2.13}$$

By imposing the orthonormality of the DCT basis e_k, the inverse discrete cosine transform (IDCT), denoted as $f[n]$, is derived as follows,

$$f[n] = \lambda_k \sqrt{\frac{2}{N}} \sum_{n=0}^{N-1} F[k] \cos\left(\frac{\pi}{N}\left(n + \frac{1}{2}\right)k\right). \tag{2.14}$$

2.4 Mean Squared Error (MSE)

The mean squared error (MSE) is one of the most popular quantitative performance metrics in the field of signal and image processing. Throughout, we consider a greyscale image which is represented by an image function $f : X \to \mathbb{R}_g$, where $X \subset \mathbb{R}^2$ and $\mathbb{R}_g$ denotes the greyscale range space that is positive and bounded, i.e., $\mathbb{R}_g = [A, B] \subset (0, \infty)$. As such, we may define the signal/image function space $\mathcal{F}(X) = \{f : X \to \mathbb{R}_g \mid f \text{ is measurable}\}$. Moreover, the consequence of the boundedness of the image function $f(x)$ is that $\mathcal{F}(X) \subset L^p(X)$ for all $p \geqslant 1$, where the $L^p(X)$ function spaces are defined in the usual way. For any $p \geqslant 1$, the L^p norm can be used to define a metric d_p on $\mathcal{F}(X)$: For $f, g \in \mathcal{F}(X)$, we have the following family p-metrics,

$$d_p(f, g) = \|f - g\|_p = \left[\int_X [f(x) - g(x)]^p dx\right]^{\frac{1}{p}}. \tag{2.15}$$

Our primary concern is the approximation of functions in the case $p = 2$, i.e., the Hilbert space, $L^2(X)$. In this case, the distance between two functions $f, g \in L^2(X)$ is given by

$$d_2(f, g) = \|f - g\|_2 = \left[\int_X [f(x) - g(x)]^2 dx\right]^{\frac{1}{2}}. \tag{2.16}$$

Notice that, technically speaking, the integrals of X presented above represent double integrals since $X \subset \mathbb{R}^2$.

With reference to Section 2.2, we recall that digital images may be represented by matrices. For any two $n_1 \times n_2$ matrices, e.g., $u = \{u_{ij}\}$ and $v = \{v_{ij}\}$, the two dimensional versions of the l_p metric defined over vectors in $\mathbb{R}^n$ are given by,

$$d_p(u, v) = \left[\sum_{i=1}^{n_1}\sum_{j=1}^{n_2} |u_{ij} - v_{ij}|^p dx\right]^{\frac{1}{p}}. \tag{2.17}$$

As a special case of the Euclidean metric, $p = 2$, we have

$$d_2(u, v) = \left[\sum_{i=1}^{n_1}\sum_{j=1}^{n_2} [u_{ij} - v_{ij}]^2 dx\right]^{\frac{1}{2}}, \tag{2.18}$$

which has been widely employed in image processing algorithms. Furthermore, the mean squared error (MSE) between the two digital images u and v is expressed as follows,

$$\text{MSE}(u, v) = \frac{1}{n_1 n_2}[d_2(u, v)]^2 = \frac{1}{n_1 n_2}\sum_{i=1}^{n_1}\sum_{j=1}^{n_2}[u_{ij} - v_{ij}]^2. \tag{2.19}$$

It then follows from the MSE that the root mean squared error (RMSE) is given by,

$$\mathrm{RMSE}(u,v) = \sqrt{\mathrm{MSE}(u,v)} = \left[\frac{1}{n_1 n_2}[d_2(u,v)]^2\right]^{1/2} = \left[\frac{1}{n_1 n_2}\sum_{i=1}^{n_1}\sum_{j=1}^{n_2}[u_{ij}-v_{ij}]^2\right]^{1/2}. \quad (2.20)$$

In fact, MSE and RMSE are useful and popular not only because they can characterize the average error per pixel, which allows one to compare errors among the pairs of images of different sizes, but also because it satisfies some desirable mathematical properties such as convexity, symmetry and differentiability [21].

2.5 Structural Similarity (SSIM)

It is well-known [21, 41] that MSE and other L^2-based distances do not measure visual quality very well. The structural similarity (SSIM) [25] was proposed to measure image fidelity in terms of structural information which is consistent with the human visual system, as opposed to the usual L^2 image distortion measurement, i.e., MSE. Over the last few years, the SSIM has been intensively applied in a wide range of academic research related to image and video processing, for instance, image compression and video streaming [22, 23, 24]. Moreover, the best approximation problem that involves SSIM-based optimization criterion has been effectively investigated and developed in [17]. In essence, the "local SSIM", which represents the SSIM value between two local image patches, denoted as x and y, measures the similarities of three elements of the patches: the similarity $l(x,y)$ of the local patch luminance or brightness, the similarity $c(x,y)$ of the local patch contrasts, and the similarity $s(x,y)$ of the local patch structures [25]. Mathematically, $l(x,y)$, $c(x,y)$ and $s(x,y)$ are defined as follows,

$$l(x,y) = \frac{2\bar{x}\bar{y}+\epsilon_1}{\bar{x}^2+\bar{y}^2+\epsilon_1}, \quad (2.21)$$

$$c(x,y) = \frac{2s_x s_y+\epsilon_2}{s_x^2+s_y^2+\epsilon_2}, \quad (2.22)$$

$$s(x,y) = \frac{s_{xy}+\epsilon_3}{s_x s_y+\epsilon_3}, \quad (2.23)$$

where ϵ_1, ϵ_2 and ϵ_3 are numerical stability coefficients which can be adjusted to keep the denominators of the above equations away from zero and which accommodate the

perception of the human visual system. With regard to Eqs. (2.21), (2.22) and (2.23), we also have that

$$\bar{x} = \frac{1}{N}\sum_{i=1}^{N} x_i, \tag{2.24}$$

$$s_x^2 = \frac{1}{N-1}\sum_{i=1}^{N}(x_i - \bar{x})^2, \tag{2.25}$$

and

$$s_{xy} = \frac{1}{N-1}\sum_{i=1}^{N}(x_i - \bar{x})(y_i - \bar{y}), \tag{2.26}$$

which denote mean values, variances and covariances of local image patches x and y, respectively. In practice, the local SSIM index between x and y is usually expressed as the product of l, c and s with appropriate powers. In the special case of $\epsilon_2 = 2\epsilon_3$, the SSIM function, $S(x, y)$, becomes a product of two terms,

$$S(x,y) = \left[\frac{2\bar{x}\bar{y} + \epsilon_1}{\bar{x}^2 + \bar{y}^2 + \epsilon_1}\right]\left[\frac{2s_{xy} + \epsilon_2}{s_x^2 + s_y^2 + \epsilon_2}\right] = S_1(x,y)S_2(x,y). \tag{2.27}$$

The SSIM function presented in the above equation is indeed the primary form that is being employed and discussed in this thesis.

Once again, let us consider $x, y \in \mathbb{R}^N$ as two N-dimensional signal or image blocks, and the SSIM function in Eq. (2.27). Notice that $-1 \leq S(x, y) \leq 1$, and $S(x, y) = 1$ implies that $x = y$. $S(x, y) = S(y, x)$ indicates that the SSIM function is symmetric with respect to the signals or image blocks. If we look closer at the luminance component, i.e., $l(x, y)$, of SSIM function, it may be rewritten as follows,

$$l(x,y) = \frac{2(\bar{y}/\bar{x}) + \epsilon_1'}{1 + (\bar{y}/\bar{x})^2 + \epsilon_1'}, \tag{2.28}$$

depending upon the ratio $\bar{y}/\bar{x}$, where $\epsilon_1' = \epsilon_1/\bar{x}^2$. Consequently, $l(x, y)$ shows the potential of accommodating Weber's model of perception. Let us consider y as an approximation of x and described by the following form,

$$y = x + r, \tag{2.29}$$

where r denotes the residual term of the approximation, i.e., y, of x. If we let $\Delta\bar{x}$ represents the mean value of the residue r, it then follows that

$$\bar{y} = \bar{x} + \Delta\bar{x}. \tag{2.30}$$

As such, rewriting Eq. (2.28) yields

$$l(x,y) = \frac{2(1+\frac{\Delta\bar{x}}{\bar{x}}) + \epsilon_1'}{1+(1+\frac{\Delta\bar{x}}{\bar{x}})^2 + \epsilon_1'}, \tag{2.31}$$

which is consistent with Weber's model of perception [25], according to Eq. (2.2). In other words, if the ratio $\Delta\bar{x}/\bar{x}$ is a constant, then for large $\bar{x}$, greater deviation in $\bar{y}$ would be assigned in order to keep the luminance measurement, i.e., $l(x,y)$, as a constant, where $l(x,y)$ is somehow relevant to the just noticeable difference with respect to luminance.

As mentioned earlier, SSIM function $S(x,y)$ is a **similarity function**: If x and y are "close", then $S(x,y)$ is close to 1. A **dissimilarity measure** may be defined for $S(x,y)$ as follows:

$$T(x,y) = 1 - S(x,y). \tag{2.32}$$

If x and y are "close", then $T(x,y)$ is close to zero. Note that $0 \leq T(x,y) \leq 2$. Finally, let us define the following zero-mean blocks,

$$x_0 = x - \bar{x}I, \quad y_0 = y - \bar{y}I \quad \Longrightarrow \quad \bar{x}_0 = \bar{y}_0 = 0, \tag{2.33}$$

where I represents the identity matrix. Then, we may obtain

$$||x_0 - y_0||^2 = \sum_{k=1}^{N}(x_{0,k} - y_{0,k})^2 = (N-1)[s_{x_0}^2 + s_{y_0}^2 - 2s_{x_0y_0}]. \tag{2.34}$$

After a little manipulation, it follows from Eq. (2.27) that

$$1 - S_2(x_0,y_0) = \frac{1}{N-1}\frac{||x_0-y_0||^2}{s_{x_0}^2 + s_{y_0}^2}, \tag{2.35}$$

As such, the function,

$$0 \leq T(x_0,y_0) \leq 2, \tag{2.36}$$

which can be viewed as a variance-weighted squared L^2 distance function [25], showing a connection between the SSIM-based approximation and the regular L^2 approximation.

Chapter 3

Intensity-based Measures in Generalized Weber's Model of Perception

In this chapter, we first review early research effort on "Weberizing" the distance between two signal or image functions u and v. The "Weberization" is incorporated with this distance function by considering a ratio of u and v. Then, we examine a novel method of producing "Weberized" distance functions, which are motivated by generalized Weber's model of perception, by employing appropriate nonuniform measures on greyscale range. In the end, we show the equivalence of the L^2 metric and the logarithmic L^2-based metric, from which we cannot generate the normed signal or image function spaces.

3.1 Previous Work on Weberizing Distance Functions

First of all, let us concisely set up some basic mathematical ingredients of our formalism [1]:

1. The **base (or pixel) space** $X \subset \mathbb{R}$ on which our signals/images are supported. Without loss of generality, we can assume that $X = [0, 1]$ for the case of signal functions. For images, $X = [0, 1]^2$. In digital image processing, X can be the set of pixel locations (i, j), $1 \leq i \leq n_1, 1 \leq j \leq n_2$.
2. The **greyscale range** $\mathbb{R}_g = [A, B] \subset (0, \infty)$.
3. The **signal/image function space** $\mathcal{F}(X) = \{u : X \to \mathbb{R}_g \mid u \text{ is measurable}\}$. Note

that from our definition of the greyscale range $\mathbb{R}_g$, $u \in \mathcal{F}(X)$ is positive and bounded, i.e., $0 < A \leq u(x) \leq B < \infty$ for almost every (a.e.) $x \in X$. A consequence of this boundedness is that $\mathcal{F}(X) \subset L^p(X)$ for all $p \geqslant 1$, where the $L^p(X)$ function spaces are defined in the usual way. For any $p \geqslant 1$, the L^p norm can be used to define a metric d_p on $\mathcal{F}(X)$: For $u, v \in \mathcal{F}(X)$, $d_p(u, v) = \|u - v\|_p$. Our primary concern is the approximation of functions in the case $p = 2$, i.e., the Hilbert space, $L^2(X)$. In this case, the distance between two functions $u, v \in L^2(X)$ is given by

$$d_2(u, v) = \|u - v\|_2 = \left[\int_X [u(x) - v(x)]^2 dx\right]^{\frac{1}{2}}. \tag{3.1}$$

In what follows, we consider $v(x)$ to be an approximation to $u(x)$ and let $\{\phi_k\}_{k=1}^{\infty}$ denote a set of basis functions (e.g., complete orthonormal basis) in $L^2[a, b]$. If $v(x)$ is a linear combination of basis functions, i.e., $v(x) = \sum_{k=1}^{N} c_k \phi_k(x)$, the usual squared L^2 error is given by

$$\Delta_N^2 = \int_a^b \left[u(x) - \sum_{k=1}^{N} c_k \phi_k(x)\right]^2 dx. \tag{3.2}$$

In the case that the $\{\phi_k\}_{k=1}^{\infty}$ form a complete orthonormal basis in the Hilbert space of functions $L^2[a, b]$, it is well-known that Δ_N^2 is minimized when

$$c_k = \langle u, \phi_k \rangle, \tag{3.3}$$

the so-called "Fourier coefficients" of u, for $1 \leq k \leq N$.

Generally, a Weberization of the L^2 metric can be achieved when we incorporate some nonnegative intensity-based weight function, denoted as $g(u(x), v(x))$, which is dependent upon one or both of the image functions $u(x)$ and $v(x)$ at each pixel, inside the L^2-based distance function in Eq. (3.2). Then, we obtain a Weberized L^2 distance function as follows [1, 2],

$$W_N^2 = \int_a^b g(u(x), v(x)) \left[u(x) - \sum_{k=1}^{N} c_k \phi_k(x)\right]^2 dx. \tag{3.4}$$

In fact, one simple and effective method of Weberizing the L^2-based metric is to divide the difference $|u(x) - v(x)|$ by $|u(x)|$ or $|v(x)|$. As a special case, when $g(u(x), v(x)) =$

$g(u(x)) = u(x)^{-2}$, Eq. (3.4) becomes

$$\begin{aligned} W_N^2 &= \int_a^b \frac{1}{u^2(x)} \left[u(x) - \sum_{k=1}^{N} c_k \phi_k(x) \right]^2 dx \\ &= \int_a^b \left[1 - \frac{\sum_{k=1}^{N} c_k \phi_k(x)}{u(x)} \right]^2 dx, \end{aligned} \tag{3.5}$$

where the function $u(x)^{-2}$ denotes the intensity-based weight function, in which regions of higher (lower) intensity values have a lower (higher) weight, which conforms to Weber's model of perception. This particular distance function involves the ratio of the reference function u and its approximation v. As examined in [1], Eq. (3.5) is convenient to work with in terms of some best approximation problems. Note, however that, strictly speaking, it is not a metric since the weighting function $g(u(x))$ is not symmetric in its argumentation. This complication was addressed in [1]. Furthermore, let us analyze the following distance function,

$$W_N^2 = \int_a^b g(u(x)) \left[u(x) - \sum_{k=1}^{N} c_k \phi_k(x) \right]^2 dx,$$

which is a general form of Eq. (3.5). To determine the optimal coefficients that minimize W_N^2 in Eq. (3.4), we impose the stationarity conditions, which is given by

$$\frac{\partial W_N^2}{\partial c_p} = -2 \int_a^b g(u(x)) \left[u(x) - \sum_{k=1}^{N} c_k \phi_k(x) \right] \phi_p(x) dx = 0, \quad 1 \leq p \leq N. \tag{3.6}$$

After some rearrangement, we obtain

$$\sum_{k=1}^{N} c_k \int_a^b g(u(x)) \phi_k(x) \phi_p(x) dx = \int_a^b g(u(x)) u(x) \phi_p(x) dx, \tag{3.7}$$

which represents a linear system of equation in the unknowns c_k having the form

$$\mathbf{Ac} = \mathbf{b}, \tag{3.8}$$

where

$$a_{kp} = \int_a^b g(u(x)) \phi_k(x) \phi_p(x), \quad b_p = \int_a^b g(u(x)) u(x) \phi_p(x), \quad 1 \leq p, k \leq N. \tag{3.9}$$

For each $p \in \{1, 2, ..., N\}$, we multiply both sides of Eq. (3.7) by c_p and sum over all p to obtain following results,

$$\int_a^b g(u(x))v^2(x)dx = \int_a^b g(u(x))u(x)v(x)dx, \tag{3.10}$$

which shows the weighted average of $v^2(x)$ is equal to weighted average of $u(x)v(x)$ when the "Weberized" L^2 error function in Eq. (3.4) is minimized. And Eq. (3.10) is named weighted average equation of "Weberized" L^2-based distance function.

3.2 Generalized Weber's Law and Nonuniform Measures on Greyscale Range Space

In this section, we introduce the method of constructing generalized Weber-based distance functions between $u(x)$ and $v(x)$ in the signal/image function space $\mathcal{F}(X)$ by using nonuniform measures that are supported on greyscale range space $\mathbb{R}_g = [A, B]$. This idea was introduced in [4] to construct the metrics to deal with image function approximation, and employed in [1] to show the logarithmic L^1 metric between two functions u and v accommodates Weber's standard model of perception. Note Weber's standard model of perception refers to Eq. (2.3), where $a = 1$.

3.2.1 Nonuniform Greyscale Measures and Logarithmic Metrics

As before, we denote the base space as $X = [a, b]$. Now, let us consider two functions u, v and define two subsets of X over the greyscale range space as follows [4],

$$X_u = \{x \in X | u(x) \leq v(x)\}, \quad X_v = \{x \in X | v(x) \leq u(x)\}. \tag{3.11}$$

Note that $X = X_u \cup X_v$, and a possible situation is sketched in the figure below [3]. In usual L^p metrics for $1 \leq p \leq \infty$, the contribution of each strip in Figure 3.1 to the integral will be an appropriate power of $|u(x) - v(x)|$, which implies a uniform weighting or measure on the intensity axis. From the nonuniform measures, however, the "distance" between such two image functions $u(x)$ and $v(x)$ is not necessarily $|u(x) - v(x)|$ but rather the sizes of the intervals $(u(x), v(x)] \subset \mathbb{R}_g$ and $(v(x), u(x)] \subset \mathbb{R}_g$, i.e., $\nu(u(x), v(x)]$ or $\nu(v(x), u(x)]$, as assigned by a measure ν that is supported on the greyscale interval $\mathbb{R}_g$. Then, the distance between u and v associated with the measure ν is defined in the following,

$$D(u, v; \nu) = \int_{X_u} \nu(u(x), v(x)]dx + \int_{X_v} \nu(v(x), u(x)]dx. \tag{3.12}$$

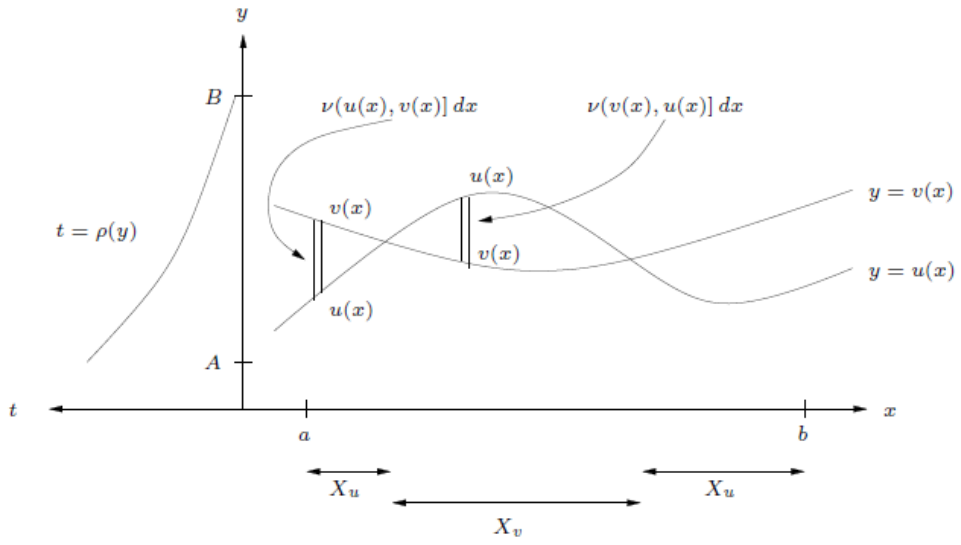


Figure 3.1: Sketch of two nonnegative greyscale functions $u(x)$ and $v(x)$ with strips of width dx that will contribute to the distance $D(u, v; \nu)$.

For convenience, we assume that the greyscale measure ν is sufficiently regular, i.e., it can be defined in terms of a continuous, nonnegative density function $\rho(y) : \mathbb{R}_g \to [0, \infty)$ as follows,

$$\nu(y_1, y_2] = \int_{y_1}^{y_2} \rho(y)\ dy = P(y_1) - P(y_2), \tag{3.13}$$

where $(y_1, y_2] \subset \mathbb{R}_g$ and $P'(y) = \rho(y)$, i.e., $P(y)$ is an antiderivative of $\rho(y)$. The distance function in Eq. (3.12) is then given by

$$\begin{aligned} D(u, v; \nu) &= \int_{X_u} \left[\int_{u(x)}^{v(x)} \rho(y) dy \right]\ dx + \int_{X_v} \left[\int_{v(x)}^{u(x)} \rho(y) dy \right]\ dx \\ &= \int_{X_u} [P(v(x)) - P(u(x))]\ dx + \int_{X_v} [P(u(x)) - P(v(x))]\ dx \\ &= \int_X |P(u(x)) - P(v(x))| dx. \end{aligned} \tag{3.14}$$

Example 3.2.1: In the special case that ν is the uniform Lebesgue measure on $\mathbb{R}_g$, to be denoted as m_g, $\rho(y) = 1$ and $P(y) = y$. From Eq. (3.14), we obtain,

$$D(u, v; m_g) = \int_{X_u} [v(x) - u(x)]\, dx + \int_{X_v} [u(x) - v(x)]\, dx = \int_X |u(x) - v(x)| dx, \quad (3.15)$$

which is the L^1 distance between u and v.

Importantly, it now remains to define what we mean by a density function conforming to Weber's model of perception [13].

Definition 3.2.1: Given $a > 0$ and the lowest possible intensity value I_0, suppose that Weber's model of perception in Eq. (2.3) holds for a particular value of $C > 0$ and for all intensities $I \geq I_0$, the density function $\rho_a(y)$ is said to conform to Weber's model if the following condition is satisfied for all $I \geq I_0$,

$$\nu_a(I, I + \Delta I) = \int_I^{I+\Delta I} \rho_a(y) dy = K, \quad (3.16)$$

where $\Delta I = CI^a$ is the minimum perceived change in intensity at I according to Eq. (2.3), and $K > 0$ is a constant.

For this chapter, we assume there exists such a density $\rho(y)$ function associated with the measure ν, which accommodates Weber's generalized model of perception, without presenting any mathematical details. The rigorous proof of the existence and uniqueness of the density function $\rho(y)$ will be given in Chapter 4.

In terms of perception, Eq. (3.16) introduces an invariance result, which may be interpreted graphically with respect to the density function $\rho_a(y)$ in Figure 3.2 [13]. For the graph, K represents the "visual accumulation" of intensity that contributes to human visual perception. Essentially, for any $a > 0$, our goal is to find density functions $\rho_a(y)$ which conform to Weber's model of perception referring to Eq. (2.3). In [1], the important result for Weber's standard model, a = 1, states the following:

Theorem 3.2.1: The measure ν on $\mathbb{R}_g$ which accommodates Weber's standard model of perception, $a = 1$ in Eq. (2.2), over the greyscale space $\mathbb{R} \subset [0, \infty)$ is, up to a normalization constant, defined by the continuous density function $\rho_1(y) = 1/y$. For any two greyscale intensities $I_1, I_2 \in \mathbb{R}_g$,

$$\int_{I_1}^{I_1+\Delta I_1} \frac{1}{y} dx = \int_{I_2}^{I_2+\Delta I_2} \frac{1}{y} dx, \quad (3.17)$$

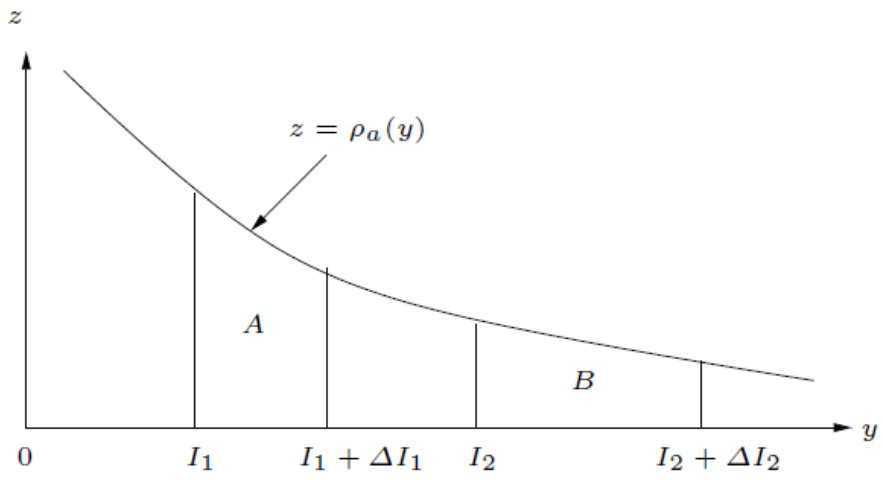


Figure 3.2: Graphical representation of the invariance result in Eq. (3.16). Area of A = Area of B.

which implies

$$\nu(I_1, I_1 + \Delta I_1) = \nu(I_2, I_2 + \Delta I_2) \tag{3.18}$$

according to Definition 3.2.1.

Example 3.2.2: In the case of $a = 1$, $\rho_1(y) = \dfrac{1}{y}$, and $P(y) = \ln y$, which is the anti-derivative of $\rho_1(y)$. From Eq. (3.14), we have

$$\begin{aligned} D(u, v; \nu) &= \int_{X_u} \left[\int_{u(x)}^{v(x)} \frac{1}{y} dy \right] dx + \int_{X_v} \left[\int_{v(x)}^{u(x)} \frac{1}{y} dy \right] dx \\ &= \int_X |\ln u(x) - \ln v(x)| \, dx \\ &= \| \ln u(x) - \ln v(x) \|_1. \end{aligned} \tag{3.19}$$

Theoretically, Eq. (3.19) is the distance function to apply in any approximation scheme which is better suited to Weber's standard model of perception. Even though the development of L^1 minimization algorithms have been steadily promoted by compressive sensing [5] and stochastic methods [6], it is still difficult to work out the best approximated function $v(x)$ by minimizing the logarithmic L^1 metric in Eq. (3.19), due to the nondifferentiability

of $D(u, v; \nu)$ with respect to the unknown coefficients in v, and computational inefficiency of the L^1 optimization algorithms. For convenience, we consider more usual and practical L^2-based analogues denoted as $d_{\log}(u, v)$, which is given by

$$\begin{aligned} d_{\log}(u, v) &= \left[\int_X \left[\log u(x) - \log v(x) \right]^2 dx \right]^{\frac{1}{2}} \\ &= \| \log u(x) - \log v(x) \|_2. \end{aligned} \tag{3.20}$$

By squaring Eq. (3.20) and expanding $v(x)$ as the linear combination of a set of basis functions $\{\phi_k\}_{k=1}^{\infty}$, we obtain the standard logarithmic L^2-based metric in the following,

$$d_{\log}^2(u, v) = \int_X \left[\log u(x) - \log(\sum_{k=1}^{N} c_k \phi_k(x)) \right]^2 dx. \tag{3.21}$$

To minimize the squared distance function in Eq. (3.21), we impose the stationarity condition on it and yield

$$\frac{\partial d_{\log}(u, v)}{\partial c_p} = -2 \int_X \left[\log u(x) - \log(\sum_{k=1}^{N} c_k \phi_k(x)) \right] \frac{\phi_p(x)}{\sum_{k=1}^{N} c_k \phi_k(x)} dx = 0, \tag{3.22}$$

where $1 \leq p \leq N$. If we let $c^*(c_1^*, c_2^*, ..., c_N^*)$ be a solution to Eq. (3.22) for any $1 \leq p \leq N$, and multiply both sides of $\frac{\partial d_{\log}(u,v)}{\partial c_p}$ by c_p^* then sum all equations, we have

$$\int_X \left[\log u(x) - \log(\sum_{k=1}^{N} c_k^* \phi_k(x)) \right] \frac{\sum_{k=1}^{N} c_k^* \phi_k(x)}{\sum_{k=1}^{N} c_k^* \phi_k(x)} dx = 0. \tag{3.23}$$

If we let $X = [a, b]$, rewriting the above equation yields

$$\frac{1}{b-a} \int_a^b \log u(x) dx = \frac{1}{b-a} \int_a^b \log(\sum_{k=1}^{N} c_k^* \phi_k(x)) dx, \tag{3.24}$$

which may be viewed as a weighted average equation of logarithmic L^2-based metric. Namely, the left hand side (LHS) of Eq. (3.24) is the mean value of $\log u(x)$ over $[a, b]$, and the right hand side (RHS) represents the mean value of $\log(\sum_{k=1}^{N} c_k^* \phi_k(x))$. Note that this type of result does not occur in the normal/standard "best L^2-based approximation" cases.

Special case 1: Constant logarithmic L^2 approximation

When $X = [a, b]$ and $v(x) = c$, which is an arbitrary constant, Eq. (3.21) then becomes

$$d^2_{\log}(u, v) = \int_a^b \left[\log u(x) - \log(c) \right]^2 dx. \tag{3.25}$$

Then, the best approximation of $u(x)$ is simply

$$c = \exp\left[\frac{1}{b-a} \int_a^b \log u(x) dx \right] \tag{3.26}$$

by solving Eq. (3.24).

In fact, Eq. (3.22) is a nonlinear system of equations, which makes the approximation problem more complicated even though we can exploit a number of effective numerical algorithms to solve it appropriately. To simplify the complexity, a modified logarithmic L^2-based distance function was introduced as follows [1, 2],

$$d^2_{\log}(u, v) = \int_X \left[\log u(x) - \sum_{k=1}^{N} c_k \phi_k(x) \right]^2 dx. \tag{3.27}$$

Now, the linear system of equations can be easily solved to yield the optimal coefficient $c^* = (c_1^*, \cdots, c_N^*)$, when we impose the stationarity condition on Eq. (3.21). As such, the best modified logarithmic L^2 approximation of $u(x)$ is given by

$$v(x) = \exp\left[\sum_{k=1}^{N} c_k^* \phi_k(x) \right]. \tag{3.28}$$

Of course, one might notice the cost of simplification is that we have lost some "Weberization" information in the approximate function (i.e., $v(x)$) over the distance function (not a metric) in Eq. (3.27).

3.2.2 Generalized Weber-based Metrics

From Theorem 3.2.1, the density function, $\rho_1(y) = 1/y$, is a decreasing function, which assigns lesser (more) weight to higher (lower) intensities as we would expect in a Weberized distance function. That being said, we would expect the density functions $\rho_a(y)$, accommodating Weber's generalized model of perception, $a > 0$ and $a \neq 1$, to be decreasing

functions of intensity y as well. Given that $\rho_1(y) = 1/y$, one might guess for any $a > 0$, the density function $\rho_a(y)$ is equal to $1/y^a$. Now, based on Definition 3.2.1 and Figure 3.2, let us prove that $\rho_a(y) \simeq 1/y^a$ is the leading order term, as $y \to \infty$, of the density function, which accommodates Weber's generalized model of perception referring to Eq. (2.3). Consider the problem: Given any intensity $I \geq I_0$ (i.e., the lowest intensity value), and an $a > 0$, find the leading-order behaviour of the density function $\rho_a(y)$ so that

$$F(I) = \int_I^{I+CI^a} \rho_a(y)dy = K, \tag{3.29}$$

where K is a constant. By assuming $\rho_b(y) = 1/y^b$, we have

$$F(I) = \int_I^{I+CI^a} \frac{1}{y^b}dy = K. \tag{3.30}$$

Differentiating both RHS and LHS of Eq. (3.30) with respect to I, we obtain

$$\frac{1 + aCI^{a-1}}{I^b(1 + CI^{a-1})^b} = \frac{1}{I^b}. \tag{3.31}$$

After multiplying I^b on both sides, the relation in Eq. (3.31) becomes

$$1 + aCI^{a-1} = (1 + CI^{a-1})^b. \tag{3.32}$$

Note that Eq. (3.32) is automatically satisfied when $a = 1$. As we show later (Chapter 4), no such density function of the form $\rho(y) = 1/y^a$ exists for $a > 1$ so here we consider only the case $0 < a < 1$. Furthermore, since Weber's law breaks down at very low intensities, we are primarily interested in the high-intensity range, i.e., the asymptotic case $x \to \infty$. In this special case, $I^{a-1} \to 0$ as $I \to \infty$. A binomial expansion of the right hand side of Eq. (3.32) yields

$$(1 + aCI^{a-1}) \approx 1 + bCI^{a-1}\,, \quad \text{as} \quad I \to \infty\,. \tag{3.33}$$

Using this result in (3.32) implies that

$$aCI^{a-1} = bCI^{a-1}\,, \tag{3.34}$$

which implies that $a = b$. Therefore, for any $0 < a < 1$, $\rho_a^1(y) = y^{-a}$ is the leading-order behaviour of density functions, which obeys the invariance condition of Weber's generalized model of perception, as $y \to \infty$. From Fig. 3.2, for any two greyscale intensities $I_1, I_2 \geq I_0$, we have

$$\int_{I_1}^{I_1+\Delta I_1} \rho_a^1(y)dy \simeq \int_{I_2}^{I_2+\Delta I_2} \rho_a^1(y)dy, \tag{3.35}$$

where $\Delta I_1 = CI_1^a$ and $\Delta I_2 = CI_2^a$ are the minimum perceived changes in intensity at I_1 and I_2, respectively [3].

Now, for $0 < a < 1$, if the leading-order approximation of the density functions, i.e., $\rho_a^1(y) = y^{-a}$, is substituted into Eq. (3.12) and (3.13), we derive the metrics between image functions produced by the nonuniform measures, which accommodates Weber's generalized model of perception, up to a multiplicative constant, as follows,

$$\begin{aligned} D(u,v;\nu) &= \int_{X_u} \left[\int_{u(x)}^{v(x)} \frac{1}{y^a} dy \right] dx + \int_{X_v} \left[\int_{v(x)}^{u(x)} \frac{1}{y^a} dy \right] dx \\ &= \int_{X_u} (v^{-a+1} - u^{-a+1})\, dx + \int_{X_v} (u^{-a+1} - v^{-a+1})\, dx \\ &= \int_X |u^{-a+1} - v^{-a+1}| dx. \end{aligned} \tag{3.36}$$

For $a = 0$, i.e., the absence of the Weber's model, Eq. (3.36) still holds, corresponding to Eq. (3.15). In this case, $\rho_0(y) = 1$, and the measure $\nu_0 = m_g$, i.e., Lebesgue measure, becomes uniform all over the greyscale range $\mathbb{R}_g$, which suggests that all greyscale intensities are weighted equally in computing the distances between image functions. In summary, along with Eq. (3.19), the metrics between image functions yielded by the measures, which are associated with the density functions, are expressed in the following,

$$\begin{aligned} a = 0: &\quad D_0(u,v;m_g) = \int_X |u(x) - v(x)| dx. \\ 0 < a < 1: &\quad D_a(u,v;\nu_a) = \int_X |u(x)^{-a+1} - v(x)^{-a+1}| dx. \\ a = 1: &\quad D_1(u,v;\nu_1) = \int_X |\ln u(x) - \ln v(x)|\, dx. \end{aligned} \tag{3.37}$$

The standard and generalized "Weberization" is essentially introduced in $D_1(u,v;\nu_1)$ and $D_a(u,v;\nu_a)$, respectively, which tolerate greater differences at higher intensity values and lower differences at lower intensity values. Moreover, as a increases, the density functions with the leading order behaviour, i.e., $\rho_a^1(y) = 1/y^a$, decrease more rapidly with respect to the intensity values y. In the meanwhile, one might wonder what the distance function $D_a(u,v;\nu_a)$, conforming to Weber's generalized model of perception, looks like for $a > 1$. In that case, we shall show the asymptotical form of density functions $\rho_a(y)$ is not $1/y^a$ but rather some other function involving the logarithm, and the relevant details are presented in Chapter 4.

3.2.3 Equivalence of Weber-based Metrics and Usual L^2 Metric

By some analysis, we discovered the distance function in Eq. (3.20) is a logarithmic L^2-based metric, but this metric does not come from a normed space. To prove the claims, we first need to show the distance function in Eq. (3.20) is a metric. At first glance, $d_{\log}(u, v)$ is a metric because we are essentially employing the logarithm functions that are continuous and bijective on usual L^2 metric. On the other hand, we can prove $d_{\log}(u, v)$ is a metric by the definition. It is evident that $d_{\log}(u, v) \geq 0$ and $d_{\log}(u, v) = d_{\log}(v, u)$ for all $u, v \in \mathcal{F}(X) \subset L^2$, where $\mathcal{F}(X)$ is the signal/image function space as before. Also, $d_{\log}(u(x), v(x)) = 0$ if and only if $u(x) = v(x)$, since $d_{\log}(u(x), v(x)) = 0 \iff \log(u(x)) = \log(v(x))$ a.e. $\iff u(x) = v(x)$ a.e.. For triangle inequality, we let

$$U(x) = \log(u(x)), \quad V(x) = \log(v(x)), \quad x \in X, \tag{3.38}$$

where X is some base or pixel space. Clearly, $U(x), V(x) \in \mathcal{F}(X) \subset L^2$ as well. Then, we have

$$d_2(U(x), V(x)) \leq d_2(U(x), W(x)) + d_2(W(x), V(x)) \tag{3.39}$$

for any $W(x) = \log(w(x)) \in \mathcal{F}(X)$, where d_2 represents the usual L^2 metric. This implies

$$d_{\log}(u(x), v(x)) \leq d_{\log}(u(x), w(x)) + d_{\log}(w(x), v(x)). \tag{3.40}$$

Hence, $d_{\log}(u, v)$ is a metric. The second part of the proof is to show we cannot define a norm which generates the metric in Eq. (3.20), i.e., $d_{\log}(u, v)$. If we assume that a norm which is associated with the logarithmic L^2 metric in Eq. (3.20) exists, then one relation between $d_{\log}(u, v)$ and this norm, denoted as $\|\cdot\|_{\log}$, is as follows,

$$\begin{aligned} \|u - v\|_{\log} &= d_{\log}(u(x), v(x)) \\ &= \left[\int_X \left[\log u(x) - \log v(x)\right]^2 dx\right]^{\frac{1}{2}}. \end{aligned} \tag{3.41}$$

Then, there exists a function $w(x) = u(x) + v(x) \in \mathcal{F}(X) \subset L^2$ such that

$$\begin{aligned} \|w - v\|_{\log} &= d_{\log}(w(x), v(x)) \\ &= \left[\int_X \left[\log w(x) - \log v(x)\right]^2 dx\right]^{\frac{1}{2}}. \end{aligned} \tag{3.42}$$

By plugging $w = u + v$ into Eq. (3.27), we obtain

$$\begin{aligned} \|u + v - v\|_{\log} &= d_{\log}(u(x) + v(x), v(x)) \\ &= \left[\int_X \left[\log(u(x) + v(x)) - \log v(x)\right]^2 dx\right]^{\frac{1}{2}} \\ &= \left[\int_X \left[\log\left(1 + \frac{u(x)}{v(x)}\right)\right]^2 dx\right]^{\frac{1}{2}}. \end{aligned} \tag{3.43}$$

The metric space indicated as $(d_{\log}(u, v), \mathcal{F}(X)$ that we are interested in is a restricted linear space, i.e., we can add and subtract elements in this space provided that they satisfy positivity condition. Thus, Eq. (3.43) becomes

$$\|u\|_{\log} = \left[\int_X \left[\log\left(1 + \frac{u(x)}{v(x)}\right)\right]^2 dx\right]^{\frac{1}{2}}. \tag{3.44}$$

which is a contradiction, since $\|u\|_{\log}$ is supposed to be a special norm of $u(x)$ without incorporating with the function $v(x)$. Therefore, we have also proved the claim that the logarithmic L^2 metric in Eq. (3.20) is not from a normed space, which completes the proof. Indeed, the reason why we care about the claim is because we aim to determine whether there is a unique optimal solution, i.e., unknown coefficients, when we minimize the distance function in Eq. (3.20). Unfortunately, the above result implies that we cannot use the theory of best approximation in normed spaces and strictly normed spaces. As such, the existence and uniqueness of the minimizer in terms of $d_{\log}(u, v)$ is not guaranteed by the usual normed space and strictly normed space. Nevertheless, we are still able to establish the existence and uniqueness theorem by some discussion which will be shown in detail in Chapter 6.

Now, suppose that our space $\mathcal{F}[a, b]$ is composed of integrable functions such that

$$1 < A \leq u(x) \leq B \tag{3.45}$$

for a.e. $x \in [a, b]$. Note we let $A > 1$ to ensure the positivity of $\log(u(x))$. Before proving the equivalence of standard L^2 and logarithmic L^2 metrics, let us recall the Mean Value Theorem:

Theorem 3.2.2: Let $g : [A, B] \to \mathbb{R}$ be continuous on $[A, B]$ and differentiable on (A, B). Then for any $y_1, y_2 \in [A, B]$, there exists a c between y_1 and y_2 such that

$$g(y_2) - g(y_1) = g'(c)(y_2 - y_1). \tag{3.46}$$

Since the logarithm function, i.e., $g(u) = \log(u)$, is continuous and differentiable on $[A, B]$, it follows, by the Mean Value Theorem, that for a.e. $x \in [a, b]$,

$$\log u(x) - \log v(x) = \frac{1}{c}(u(x) - v(x)), \tag{3.47}$$

where c lies between $u(x)$ and $v(x)$. Since $A < c < B$, we arrive at the result,

$$\frac{1}{B}d_2(u, v) \leq d_{\log}(u, v) \leq \frac{1}{A}d_2(u, v). \tag{3.48}$$

Notice that the equality in Eq. (3.48) is justified in the special case that $u(x) = v(x)$. Therefore, we show standard logarithmic L^2-based metric is equivalent to L^2 metric, i.e., the convergence in standard logarithmic L^2 metric implies the convergence in L^2 metric.

Indeed, for any fixed $a \in [0, 1)$, the generalized Weber-based L^2 metrics are given by

$$d_a(u, v; \nu_a) = \left[\int_X [u(x)^{-a+1} - v(x)^{-a+1}]^2 dx \right]^{\frac{1}{2}}, \tag{3.49}$$

where $u, v \in F(X) \subseteq L^2 \subset \mathbb{R}_g$, $A \leq u(x) \leq B$ for a.e. $x \in X = [\tilde{a}, \tilde{b}]$. Note that $\tilde{a}$ (or $\tilde{b}$) is a nonnegative constant, which is distinct from Weber's component a. By Theorem 3.2.2, we shall show the generalized Weber-based L^2 metrics are equivalent to the usual L^2 metric. Imposing the Mean Value Theorem on the antiderivatives of the density functions, i.e., $P_a(y) = y^{-a+1}$ on $[A, B]$ with $A > 0$, yields

$$u(x)^{-a+1} - v(x)^{-a+1} = \frac{1-a}{c^a}(u(x) - v(x)), \quad \text{for a.e. } x \in [\tilde{a}, \tilde{b}], \tag{3.50}$$

where c lies between $u(x)$ and $v(x)$. Since $A < c < B$, we can derive the following inequalities for a.e. $x \in [\tilde{a}, \tilde{b}]$,

$$\frac{1-a}{B^a}|u(x) - v(x)| \leq |u(x)^{-a+1} - v(x)^{-a+1}| \leq \frac{1-a}{A^a}|u(x) - v(x)|. \tag{3.51}$$

By integrating x over $[\tilde{a}, \tilde{b}]$, we find that

$$\frac{1-a}{B^a}d_2(u, v) \leq d_a(u, v; \nu_a) \leq \frac{1-a}{A^a}d_2(u, v), \tag{3.52}$$

where $0 < a < 1$, and d_2 denotes the usual L^2 norm. This implies the equivalence of L^2 and generalized Weber-based L^2 metrics. Similarly, we claim that the usual L^p metrics are equivalent to the generalized Weber-based L^p metrics, where $1 \leq p \leq \infty$. The proof of the claim, simply following the above process described by Eqs. (3.49-3.51), is omitted here.

3.3 Function Approximation using Generalized Weber-based Metrics

3.3.1 Introduction

Now, we let two nonnegative greyscale functions $u(x)$ and $v(x)$ be a reference function and an approximation to $u(x)$, respectively. Given the set $\{\phi_k\}_{k=1}^N$ that is linearly independent or orthogonal over the base space X, the approximation $v(x)$ is normally expressed as the linear combination of N-basis, i.e., $\phi_1, \cdots, \phi_N$, in the following,

$$v(x) = \sum_{k=1}^{N} c_k \phi_k(x). \tag{3.53}$$

Given an image function $u(x) \in \mathcal{F}(X)$ on X, where $\mathcal{F}(X)$ is defined in the following,

$$\mathcal{F}(X) = \{u : X \to \mathbb{R}_g \mid u \text{ is measurable}\}, \tag{3.54}$$

our goal is to search for the best approximation of $u(x)$ in the metric space $(\mathcal{F}(X), D_a(u, v; \nu_a))$ for any $0 < a < 1$. To obtain the optimal solution, i.e., the best approximation denoted as $v(x)$, which corresponds to Eq. (3.53), we shall minimize the following distance function,

$$D_a(u, v; \nu_a) = \int_X |u(x)^{-a+1} - (\sum_{k=1}^{N} c_k \phi_k(x))^{-a+1}| dx. \tag{3.55}$$

However, due to the difficulties of differentiating the integrand of $D_a(u, v; \nu_a))$ in Eq. (3.54) with respect to each coefficient c_k, and the expensive computations to achieve L^1-minimization, we typically consider the squared L^2 analogues of Eq. (3.54), which are given by

$$D_a(u, v; \nu_a) = \int_X [u(x)^{-a+1} - (\sum_{k=1}^{N} c_k \phi_k(x))^{-a+1}]^2 dx. \tag{3.56}$$

After imposing stationarity conditions on Eq. (3.55), for any $1 \leq p \leq N$, we have

$$\int_X [u(x)^{-a+1} - (\sum_{k=1}^{N} c_k \phi_k(x))^{-a+1}] \frac{\phi_p(x)}{(\sum_{k=1}^{N} c_k \phi_k(x))^a} dx = 0, \tag{3.57}$$

which yields an extremely complicated nonlinear system of equations in terms of the unknown coefficients $c_1, \cdots, c_N$. In order to obtain the optimal coefficients to minimize the

distance function in Eq. (3.56), one may employ some optimization algorithms such as gradient descent method, Gauss-Newton method and so on. Indeed, some of those numerical schemes were exploited in the following best approximation examples in both 1 dimensional and 2 dimensional cases. More importantly, we shall visualize the properties of Weber's (generalized) model of perception, i.e., $0 < a \leq 1$, by comparing the best (generalized) Weber-based approximation with the usual best L^2 approximation.

3.3.2 Some Interesting Examples in Range-based Approximations

Example 1: Consider the following step function on $X = [0, 1]$,

$$u(x) = \begin{cases} 2, & 0 \leq x \leq \frac{1}{2} \\ 4, & \frac{1}{2} < x \leq 1. \end{cases} \tag{3.58}$$

In Figure 3.3, we have shown the plots which are the best approximations for the cases $a = 0$ (best L^2), $a = 0.25, 0.5$ (generalized Weber) and $a = 1$ (standard Weber) using the following DCT (discrete cosine transform) basis functions,

$$\phi_k(x) = \lambda_k \sqrt{\frac{2}{N}} \cos\left(\frac{\pi}{N}\left(x + \frac{1}{2}\right)k\right), \tag{3.59}$$

where $k = 0, 1, \cdots, 4$, and

$$\lambda_k = \begin{cases} \frac{1}{\sqrt{2}}, & k = 0 \\ 1, & k \neq 0, \end{cases} \tag{3.60}$$

denoting the normalization constants for different values of k. In the experiment, we have discretized the function $u(x)$ in Eq. (3.58) into an N-vector to represent a digital signal, which is produced by sampling the function $u(x)$ at N equidistant points between 0 and 1. In this case, we let $N = 512$, and Eq. (3.58) becomes

$$\phi_k(n) = \lambda_k \sqrt{\frac{2}{N}} \cos\left(\frac{\pi}{N}\left(n + \frac{1}{2}\right)k\right), \tag{3.61}$$

where $n = 0, 1, \cdots, 512$.

From Figure 3.3, we notice that the approximations for $a = 0.25$, $a = 0.5$ and $a = 1$

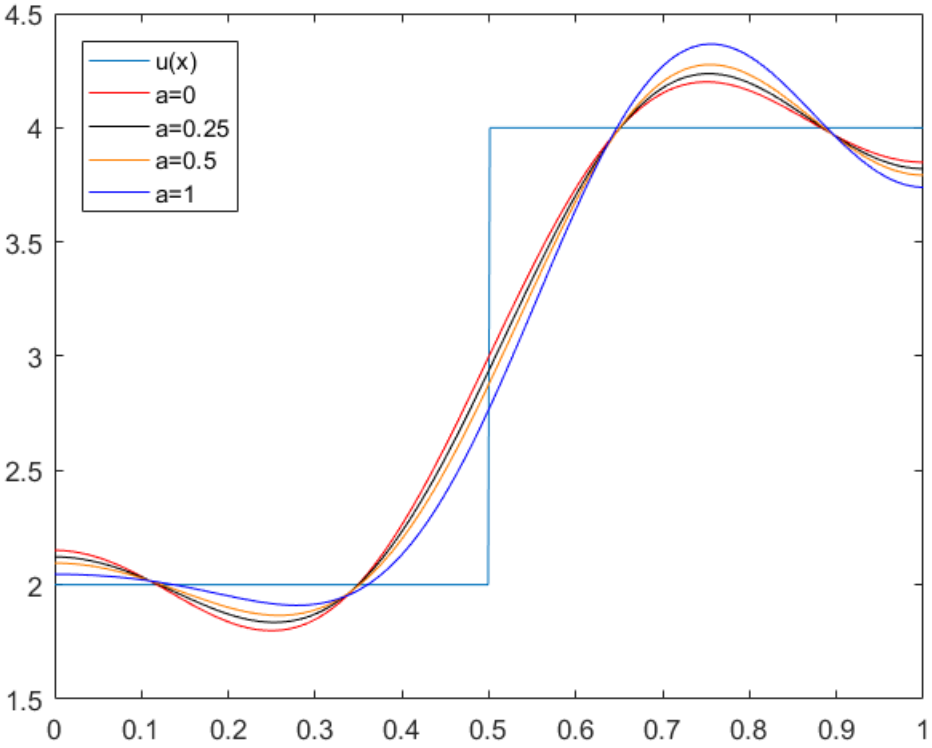


Figure 3.3: Best L^2, standard Weberized L^2 ($a = 1$), generalized Weber-based L^2 ($a = 0.25, 0.5$) approximations to the step function using $N = 5$ DCT (discrete cosine transform) basis functions.

show less variance from $u(x)$ than the best L^2 approximation ($a = 0$) over the lower intensities ($0 \leq x \leq 1/2$) and higher variance over the higher intensities ($1/2 < x \leq 1$), which accommodates (generalized) Weber's model of perception. Moreover, the difference between the best approximations and $u(x)$, i.e., reference function, decreases (increases) as $a \in [0, 1]$ increases in the lower (higher) intensity regions. Because when a increases, the associated density function $\rho_a(y) = 1/y^a$ is decreasing more rapidly, which results in more deviation over the higher intensity regions.

Example 2: Let us consider a 512×512-pixel 8 bpp image, which is composed by four squares with greyscale values 60, 128, 128 and 220. Figure 3.4 shows original square image,

the best L^2 approximations ($a = 0$) and standard Weberized L^2 ($a = 1$) using the following 2 dimensional DCT basis,

$$\Phi_{kl}(n, m) = \phi_k(n)\phi_l(m), \tag{3.62}$$

where $\phi_k(n)$ was defined in Eq. (3.61), and $0 \leq k, l \leq 14$. We expect standard Weberized L^2 approximation exhibits greater/lesser deviation than the best L^2 approximation at higher/lower greyscale levels.

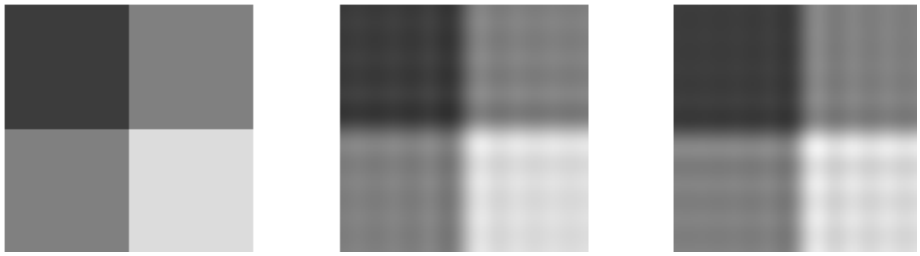


Figure 3.4: **Left and right, respectively:** Original square image, best L^2 ($a = 0$) and best standard Weberized ($a = 1$) L^2 approximations to the squares image (512×512) using 15×15 2 dimensional DCT basis

By inspection, we realize there is less/more texture in the black/white square in the right image comparing with the black/white square in the middle image. As such, the black square in the best standard Weberized L^2 approximation and white square in the best usual L^2 approximation are treated as the "better" approximation to the original image in terms of human visual perception. Essentially, the results described above correspond to our expectation that the best "Weberized" L^2 method ($a = 1$) produces a better/worse approximation than best L^2 approximation method ($a = 0$) over the lower/higher intensity regions.

Example 3: Consider a 512×512-pixel 8 bpp image called *Lena*. Figure 3.5 exhibits the best approximations for $a = 0$ (best L^2), $a = 0.5$ (generalized Weber) and $a = 1$ (standard Weber) using only $2 \times 2 = 4$ $2D$-DCT basis functions for each 8×8 block. Note that there are totally 4096 8×8 blocks in *Lena* image. For each block, we minimize the distance functions in Eq. (3.2), Eq. (3.56) and Eq. (3.21) in order to find the best L^2, generalized Weber-based L^2 and standard Weber-based L^2 approximations. After processing all blocks by minimizing these metrics, we can construct the associated approximation to the whole *Lena* image. Theoretically, we can perfectly reconstruct the *Lena* by every approximation method if 8×8 2D DCT basis is employed for each 8×8 block.

(a) Original image

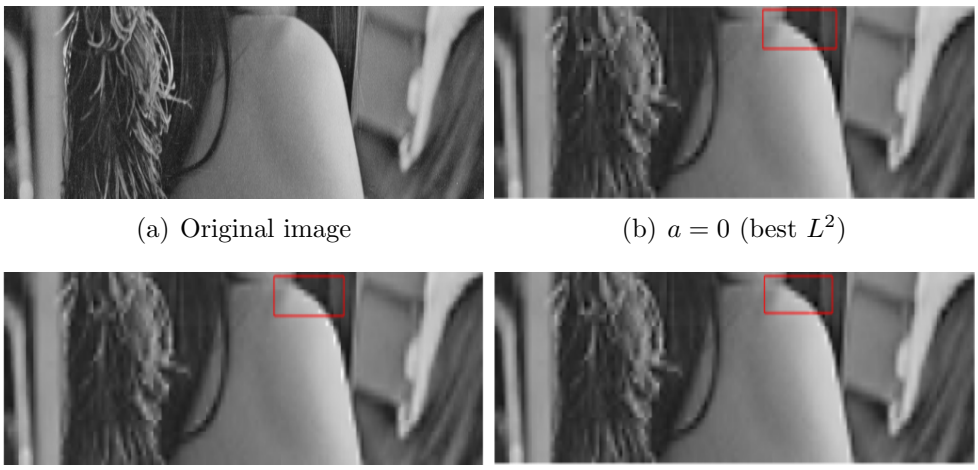
(b) $a = 0$ (best L^2)

(c) $a = 0.5$ (generalized Weber)

(d) $a = 1$ (standard Weber)

Figure 3.5: Original image *Lena* and best approximations for $a = 0$ (best L^2), $a = 0.5$ (generalized Weber) and $a = 1$ (standard Weber) using only $2 \times 2 = 4$ $2D$-DCT basis functions for each 8×8 block.

In [1], the authors state the visual effect that we expect is somewhere between the usual L^2 ($a = 0$) and standard Weberized L^2 ($a = 1$) approximations. As such, we aim to figure out the effect of generalized Weber-based approximation. First of all, it is interesting to see a number of visible "ringing artifacts" in all the approximations especially over the



(a) Original image (b) $a = 0$ (best L^2)

(c) $a = 0.5$ (generalized Weber) (d) $a = 1$ (standard Weber)

Figure 3.6: Best approximations for $a = 0$ (best L^2), $a = 0.5$ (generalized Weber) and $a = 1$ (standard Weber) using $2 \times 2 = 4$ $2D$-DCT basis over 8×8 blocks comprising the shoulder region of *Lena* image (zoomed in).

shoulder region in Figure 3.5 and 3.6. The ringing artifacts in the approximate images are caused by the distortion over regions with edges separating high and low greyscale intensities. By observing Figure 3.5, we can hardly justify the overall differences among the approximations (b), (c) and (d). However, if we look at Figure 3.6 carefully, the ringing artifacts over the shoulder region in each approximation are slightly different from each other; in particular, there are more "spikes" near edges over the region circled by red rectangle in best L^2 approximation ($a = 0$) in Figure 3.6 (b), comparing with the best standard ($a = 1$) and generalized ($a = 0.5$) Weber-based approximations in Figure 3.6 (c) and (d). This, once again, suggests the best standard and generalized Weber-based tolerate lesser differences over regions in which they assume lower intensity values.

Chapter 4

Existence, Uniqueness and Asymptotic Behaviour of Intensity-based Measures

During the process of deriving the metrics conforming to generalized Weber's model of perception, the density functions play an important role in constructing the nonuniform measures in greyscale range space to describe the distance functions between two functions u and v. However, the effectiveness of the generalized Weber-based metrics, in particular logarithmic L^2 metric, are mathematically based on the existence of the associated density functions $\rho_a(y)$ defined on the greyscale range space $\mathbb{R}_g = [A, \infty]$, where $a, A > 0$. In this chapter, we present a formal mathematical proof of the existence and uniqueness of continuous density functions $\rho_a(y)$ by Abel's equation, a well-known functional equation, which is a powerful tool to solve the embedding problem [14]. As standard Weber's law breaks down at very low and high luminance or contrast, we are also interested in showing the asymptotic behaviour of the density functions $\rho_a(y)$, satisfying invariance condition (i.e., Eq. (3.16)) illustrated in the previous chapter, for y goes to zero and infinity. These asymptotic results of density functions in essence complete the generalized version of Weber's law. Additionally, we provide a scheme that computes any number of terms of the asymptotic expansion of $\rho_a(y)$ for $y \to \infty$ and $0 < a < 1$.

4.1 Existence and Uniqueness of Greyscale Density Functions for Generalized Weber's Model

For any $a > 0$, the generalized Weber's model of perception in Eq. (2.3) suggests that the density function $\rho_a(y)$ should be a decreasing function of intensity y in greyscale range space. As such, the HVS will allow greater deviations between $u(x)$ and $v(x)$ before being perceived, with increasing intensity value. Indeed, the density function $\rho_1(y) = 1/y$ defines the logarithmic measure ν such that the L^1-based distance function between u and v becomes Eq. (3.19). Therefore, our next step is to find the density functions that conform to Weber's generalized model. However, for any $a > 0$, one might be curious about whether there exists a unique density function that accommodates Weber's generalized model of perception. From the invariance property defined in Definition 3.2.1, we then mathematically state the following existence-uniqueness theorem of the density functions conforming to Weber's generalized model.

Theorem 4.1.1: For a given $a > 0$ and a cutoff intensity value I_0, there exists a unique, continuous function $\rho_a(y)$ defined on $[I_0, \infty)$ which conforms to Weber's generalized model of perception, i.e., for any two greyscale intensities $I_1, I_2 \geq I_0$ and a particular value of $C > 0$,

$$\int_{I_1}^{I_1+\Delta I_1} \rho_a(y)dx = \int_{I_2}^{I_2+\Delta I_2} \rho_a(y)dx, \tag{4.1}$$

where

$$\Delta I_1 = CI_1^a, \quad \Delta I_2 = CI_2^a. \tag{4.2}$$

Before proceeding with the proof of the above theorem, we make two comments:

1. In the special case $a = 1$, the density function $\rho_1(y) = 1/y$ exactly, as per Theorem 3.2.1. It can easily be shown that both sides of Eq. (4.1) are equal to $\ln(1 + C)$.

2. Theorem 4.1.1 above can actually be extended to the case $a = 0$, which corresponds to an absence of Weber's law. In this case, $\rho_0(y) = 1$ which corresponds to uniform Lebesgue measure on $\mathbb{R}_g$.

We now proceed to prove Theorem 4.1.1. From elementary calculus, let us first define the antiderivative of the density functions g_a as follows,

$$P(x) = \int_{x_0}^{x} g_a(y)dy, \quad x \geq x_0 > 0, \tag{4.3}$$

where x_0 is the cutoff greyscale value. Clearly, $P(x_0) = 0$. According to the equal area condition in Eq. (3.16) and Theorem 4.1.1, we have

$$P(x + Cx^a) - P(x) = K, \quad \text{constant}, \quad \text{for} \quad x \geq x_0, \tag{4.4}$$

where $C > 0$ is the fixed Weber's constant. If we let $f(x) = x + Cx^a$, Eq. (4.4) becomes,

$$P(f(x)) - P(x) = K. \tag{4.5}$$

Dividing both sides of Eq. (4.5) by $K > 0$ gives,

$$Q(f(x)) = Q(x) + 1, \tag{4.6}$$

where $Q(x) = \frac{1}{K}P(x)$. Indeed, Eq. (4.6) is known as Abel's equation, which is a well-known functional equation that controls the iteration of $f(x)$ [15]. For instance, replacing x with $f(x)$ in Eq. (4.6) yields

$$Q(f^2(x)) = Q(f(x)) + 1 = Q(x) + 2, \tag{4.7}$$

where $f^2(x)$ denotes $f(f(x))$. Consequently, we derive the following functional equation by encoding a recurrent procedure,

$$Q(f^n(x)) = Q(x) + n, \quad n \geq 1. \tag{4.8}$$

Note that $f^n(x)$ represents the n^{th} iteration of f, i.e., $f^n(x) = f(f(\cdots f(x)))$. Now, if we multiply both sides of Eq. Eq. (4.8) by K, we have

$$P(f^n(x)) = P(x) + nK, \quad n \geq 1, \tag{4.9}$$

which can be regarded as a "unnormalized" Abel's equation for $P(x)$ in our problem of generalized Weber's law. By assuming x_0 is the initial point and defining $x_n = f^n(x_0)$, according to Eq. (4.3), Eq. (4.9) simply states that each iteration of f represents the addition of another block of area K to the integral. However, iteration of f in Eq. (4.8) counts the number of blocks instead of their areas.

Given the function $f(x)$, Abel's equation in Eq. (4.6) is said to be **solvable** if there exists a continuous solution $Q(x)$ such that Eq. (4.6) is satisfied for all $x \in X$, where $X \subset \mathbb{R}$. In what follows, we shall discuss the existence of a continuous solution of Eq.

(4.6) by imposing specific properties of $f(x)$. In particular, these properties are connected with the dynamics of iteration of f as presented below. First, let us state following theorem [15] which is very useful for the analysis of solvability of $Q(x)$ in Eq. (4.6).

Theorem 4.1.2: Abel's equation in Eq. (4.6), for Q and f real-valued functions, has a continuous solution if and only if f has no periodic orbits, in which f is said to be **acyclic**.

From the above theorem, for a continuous solution of $Q(x)$ to exist, the function $f(x)$ cannot have a periodic point. In fact, if $p \in [x_0, \infty)$ is a periodic point of $f(x)$, we have

$$f^n(p) = p, \tag{4.10}$$

for some $n \geq 1$. Note that a fixed point $f(p) = p$ is a particular example of a periodic point. Then from Eq. (4.8), for $x = p$, we obtain,

$$Q(p) = Q(p) + n, \tag{4.11}$$

which has no solution for $n \geq 1$, implying that $Q(p)$ is undefined. This is actually the part of the proof of Theorem 4.1.2.

In the case of $a = 0$, the function $f(x) = x + Cx^0 = x + C$, which represents a simple translation map. Since $C > 0$, it follows that the function, $f(x) = x + C$, has no periodic points on $[x_0, \infty)$. Hence, there exists a continuous solution of $Q(x)$ for Eq. (4.6) on $[x_0, \infty)$. For $a > 0$, the function $f(x) = x + Cx^a$ in our generalized Weber's model has only $x = 0$ as a periodic point, but $x = 0 \notin [x_0, \infty)$; thus, there exists the continuous solution for our problem over the interval $[x_0, \infty)$.

On the other hand, we can show the existence of continuous solution of $Q(x)$ for Abel's equation in Eq. (4.6) by examining the iteration dynamics of the function $f(x)$ on the interval $[x_0, \infty)$. It follows from $f(x) = x + Cx^a$ that

$$f'(x) = 1 + aCx^{a-1} > 0, \quad x \geq x_0 > 0, \tag{4.12}$$

which suggests that $f(x)$ is an monotonically increasing function on $[x_0, \infty)$. Note that

$$f(x) = x + Cx^a > x, \quad x \geq x_0, \tag{4.13}$$

implying that the graph of $f(x)$ lies above the line $y = x$ for all $x \geq x_0$. Now, if we replace x with $f(x)$, we have,

$$f^2(x) = f(x) + Cf(x)^a > f(x), \tag{4.14}$$

for all $x \geq x_0$. Continuing this procedure indefinitely will yield,

$$f^{n+1}(x) > f^n(x), \tag{4.15}$$

for all $n \geq 1$ and $x \in [x_0, \infty)$. As a consequence, we may set up the sequence of iteration functions $\{f^n(x)\}_{n=1}^{\infty}$, which is called **forward orbit** of x under iteration by $f(x)$. Clearly, the sequence $\{f^n(x)\}_{n=1}^{\infty}$ is a **strictly increasing sequence**. In addition, rewriting Eq. (4.14) gives that

$$f^2(x) = x + Cx^a + C(x + Cx^a)^a > x + 2Cx^a. \tag{4.16}$$

By replacing x with $f(x)$ in Eq. (4.16), we obtain,

$$\begin{aligned} f^3(x) &> f(x) + 2Cf(x)^a \\ &> x + Cx^a + 2Cx^a \\ &= x + 3Cx^a. \end{aligned} \tag{4.17}$$

Hence, it is straightforward to show that

$$f^n(x) > x + nCx^a, \tag{4.18}$$

by inductive arguments, for all $C > 0$, $a \geq 0$ and $x \in [x_0, \infty)$. As $n \to \infty$, the sequence of iterations $f^n(x) \to \infty$ for any $x \in [x_0, \infty)$, which shows that the sequence $\{f^n(x)\}_{n=1}^{\infty}$ is **unbounded**. This result is consistent with the fact that $f(x)$ has no periodic points. Since $f(x) = x + Cx^a$ on $[x_0, \infty)$, we can define the following mapping,

$$f : [x_0, \infty) \to [x_1, \infty), \tag{4.19}$$

where $x_1 = x_0 + Cx_0^a$. More generally, let us define the increasing sequence of points $\{x_n\}_{n=0}^{\infty}$ as follows,

$$x_{n+1} = f(x_n) = x_n + Cx_n^a, \tag{4.20}$$

and

$$I_n = [x_n, x_{n+1}], \tag{4.21}$$

for all $n \geq 0$. Then, we have

$$f : I_n \to I_{n+1}, \tag{4.22}$$

and

$$f^{-1} : I_{n+1} \to I_n, \quad n \geq 0. \tag{4.23}$$

From these properties it follows that for $n \geq 0$,

$$x \in I_n \implies f^{-n}(x) \in I_0. \tag{4.24}$$

That being said, for a fixed $n \geq 1$, given any point $x \in I_n$, we can eventually map it into I_0 after a finite number of iterations of f^{-1}. Note that $f(x)$ is an increasing one-to-one function on $[x_0, \infty)$, which implies the existence of its inverse function $f^{-1}(x)$ for all

$x \in [x_1, \infty)$. Furthermore, points in I_0 cannot be "iterated backward" since they do not belong to the range of f.

In [16], Belitskii and Lyubich discussed the existence of the solution of Abel's equation by using the dynamics of f. From several parts of the paper, we extract information which is helpful with showing that there exist solutions to Eq. (4.6) for our generalized Weber models, based on the dynamical properties of f presented above. Following the notation in [16], we let $X \subset \mathbb{R}$ denote the interval over which we consider the solution to Abel's equation in Eq. (4.6).

Proposition 4.1.1: If X is compact, Abel's equation in Eq. (4.6) has no solution, i.e., no continuous solution.

Proof: It is clear that the continuous solution $Q(x)$ on a compact set $X \subset \mathbb{R}$ is bounded. However, for any $x \in X$, the iterates $f^n(x) \in X$. Then, according to Eq. (4.8), $Q(f^n(x)) \to \infty$ as $n \to \infty$, which is a contradiction.

In consequence, we introduce the corollary as showed below.

Corollary 4.1.1: If there exists an f-invariant compact nonempty set $K \subset X$, i.e., a set K such that $f(K) \subseteq K$, then a continuous solution to Eq. (4.8) cannot exist.

Definition 4.1.1: A closed set $A \subset X$ is said to be an **absorber** if $f(A) \subset A$, i.e., A is f-invariant, and for any point $x_0 \in X$, there exists a neighbourhood V_0 of x_0 and a number $k_0 = k(x_0)$ such that $f^n(V_0) \subset A$ for $n \geq k_0$.

Definition 4.1.2: A closed set $N \subset X$ is said to be a **nozzle** if its preimage, defined as follows,

$$F^{-1}(N) = \{x \in X \mid f(x) \in N\}, \tag{4.25}$$

is a subset of N, i.e., $F^{-1}(N) \subset N$, and for any $x_0 \in X$, there exists a neighbourhood W_0 of x_0 and a number $l_0 = l(x_0)$ such that $f^{-n}(W_0) \subset N$ for $n \geq l_0$.

From Definitions 4.1.1 and 4.1.2, it is also worth to mention that the set $X \setminus N$ is invariant if N is a nozzle for f; if f is a homeomorphism, i.e., f is continuous and bijective on X and f^{-1} is continuous, then a nozzle for f is an absorber for f^{-1} [16].

Theorem 4.1.3: If Abel's equation in Eq. (4.6) is solvable, i.e., has a continuous solution on X, then there exist an absorber A and a nozzle N with empty intersection. Conversely, let f be a homeomorphism and suppose that the space X is normal, i.e., any two disjoint sets $A, B \subset X$ have nonintersecting open neighbourhoods. If there exist an absorber and a nozzle with empty intersection, then Eq. (4.6) is solvable.

For a fixed $C > 0$ and $a \geq 0$, we now employ Theorem 4.1.3 on testing Abel's equation in Eq. (4.6) with $f_a(x) = x + Cx^a$ on the interval $X = [x_0, \infty)$. From our earlier discussion on the iteration dynamics of $f_a(x)$, we may conclude the following: For any $n \geq 1$, let

$$N = \bigcup_{i=0}^{n-1} I_i = [x_0, x_n] \tag{4.26}$$

and

$$A = \bigcup_{i=n+1}^{\infty} I_i = [x_{n+1}, \infty), \tag{4.27}$$

where x_n and I_i corresponds to Eq. (4.20) and Eq. (4.21), respectively. Apparently, A is an absorber and N is a nozzle for f_a. Note that $A \cap N = \emptyset$. Additionally, $f_a(x)$ are homeomorphic functions on $[x_0, \infty)$. Because for any $a \geq 0$, the function $f_a(x) = x + Cx^a$ is bijective and continuous on $[x_0, \infty)$, and f_a and f_a^{-1} map any open set $I = (c, d)$ to basic open sets, where $c, d \in [x_0, \infty)$. From Theorem 4.1.3, Abel's equation in Eq. (4.6) is solvable, which implies that there exists a unique continuous solution $Q(x)$ on $X = [x_0, \infty)$. Since $P(x) = KQ(x)$, there exists a unique function $P(x)$, as defined in Eq. (4.3), which satisfies the equal-area condition in Eq. (4.4). However, it still remains to solve the problem of determining whether the integral function $P(x)$ is determined by a unique density function $g(x)$ and whether $g(x)$ is continuous.

From Eq. (4.4), we can define the following integral

$$G(x) = \int_x^{x+Cx^a} g_a(y)dy = K, \tag{4.28}$$

where K is a constant. By assuming that $g_a(y)$ is a continuous function, and differentiating both sides of Eq. (4.28) with respect to x, we obtain

$$g_a(x + Cx^a)(1 + aCx^{a-1}) - g_a(x) = 0. \tag{4.29}$$

Again, defining the cutoff greyscale value $x_0 > 0$ and rewriting Eq. (4.29) yields,

$$g_a(f(x))f'(x) - g_a(x) = 0, \tag{4.30}$$

where

$$f(x) = x + Cx^a, \quad x \in [x_0, \infty). \tag{4.31}$$

For $a \geq 0$ and $C > 0$, given a generalized Weber's model that defines the function $f(x)$ in (4.31), our goal is to impose previous result that there exists a unique solution of integral

function $P(x)$ to establish the existence and uniqueness of a continuous solution $g(x)$ to Eq. (4.30) over the interval $X = [x_0, \infty)$. If such a solution exists, we then integrate both sides of this equation as follows,

$$\int_{x_0}^{x} \Big[g_a(f(y))f'(y) - g_a(y) \Big] dy = 0. \tag{4.32}$$

By the change of variables and the Chain Rule, it follows that

$$\int_{f(x_0)}^{f(x)} g_a(t)dt - \int_{x_0}^{x} g_a(y)dy = 0, \tag{4.33}$$

which is equivalent to

$$\int_{f(x_0)}^{x_0} g_a(y)dy + \int_{x_0}^{f(x)} g_a(y)dy - \int_{x_0}^{x} g_a(y)dy = 0. \tag{4.34}$$

Hence, we obtain,

$$\int_{x_0}^{f(x)} g_a(y)dy - \int_{x_0}^{x} g_a(y)dy = \int_{x_0}^{f(x_0)} g_a(y)dy = K. \tag{4.35}$$

Now, if we define $P(x)$ as in Eq. (4.3), then the above equation is equivalent to Eq. (4.5). As such, $g_a(x)$ is a density function which defines the integral $P(x)$ which satisfies the Weber equal-area condition.

Indeed, Eq. (4.30) is a special case of the family of linear functional equations having the form,

$$P(x)\psi(F(x)) + Q(x)\psi(x) = \gamma(x), \tag{4.36}$$

and studied by Belitskii and Lyubich [16]. From Theorem 1.1 in [16], let us state the following Theorem in terms of functional equation in (4.36) which is defined for all $x \in X$, where X is a topological space.

Theorem 4.1.4: Let $F : X \to X$ be a homeomorphism. Then if the Abel equation associated with Eq. (4.36),

$$\phi(F(x)) - \phi(x) = 1, \tag{4.37}$$

is solvable, i.e., has a continuous solution ϕ, then Eq. (4.36) is totally solvable, i.e., it is solvable for every continuous function $\gamma(x)$ on X.

The problem is that the mapping $f(x) = x + Cx^a$ is not a homeomorphism on our space

$X = [x_0, \infty)$ for all $C > 0$ and $a \geq 0$. Indeed, $f(x)$ is not a surjective function on X, since for any $y \in [x_0, x_1)$, there is no element $x \in X$ such that $f(x) = y$, where $x_1 = x_0 + Cx_0^a$. Nevertheless, $f(x)$ is injective on X, which provides another possibility to make the above theorem effectively imposed. Of course, the price paid by the fact that $f(x)$ is just injective is that the topological space X has to satisfy more additional conditions, i.e., X is locally compact and countable at infinity (l.c.c.i), referring to Corollary 1.6 of [16].

Definition 4.1.3: A topological space X is said to be locally compact if every point $x \in X$ has a compact neighbourhood, i.e., there exists an open set U and a compact set K such that $x \in U \subseteq K$.

Definition 4.1.4: [16] A topological space X is said to be countable at infinity if there exists a covering

$$X = \bigcup_{i=1}^{\infty} K_i, \tag{4.38}$$

where the K_i are compact. (As stated in [16], without loss of generality, one can assume that $K_i \subset K_{i+1}$.)

Now, the first concern is whether the space $X = [x_0, \infty)$ is locally compact or not. Technically speaking, the point x_0 cannot have open neighbourhoods of the form $(x_0 - \delta, x_0 + \delta)$ which lie in X, where $\delta > 0$. To address this issue, we may consider the base space as $X = (x_0, \infty)$, which is clearly l.c.c.i, i.e., locally compact and countable at infinity. By previous analysis, we know the Abel equation in Eq. (4.37) is solvable, where $F = f : I_n \to I_{n+1}$, according to Eq. (4.19). Consequently, it is straightforward to see that the generalized functional equations in Eq. (4.36) are totally solvable, which implies the solvability of the special case of Eq. (4.30). In conclusion, for any $a \geq 0$ and $x_0 > 0$, there exists a unique density function $g_a(x)$, satisfying the equal area condition with respect to generalized Weber's model, on the interval $[x_0, \infty]$. This completes the proof of Theorem 4.1.1.

4.2 More Detailed Asymptotic Behaviour of Greyscale Density Functions

4.2.1 Asymptotic analysis of density functions for the case $0 < a < 1$

In Section 3.2.2, we have proved for any $0 < a < 1$, $\rho_a(y) = y^{-a}$ is the leading-order behaviour of density functions that accommodate generalized Weber's model of perception, as $y \to \infty$. As a result, we obtain the following theorem [3].

Theorem 4.2.1: For $0 < a < 1$, the density function $\rho_a(y)$ which accommodates generalized Weber's model of perception is, to leading order, $\rho_a(y) = 1/y^a$ as $y \to \infty$: For any two greyscale intensities $I_1, I_2 \geq I_0 > 0$,

$$\int_{I_1}^{I_1+\Delta I_1} \frac{1}{y^a} dx \simeq \int_{I_2}^{I_2+\Delta I_2} \frac{1}{y^a} dx, \tag{4.39}$$

where I_0 represents the lowest greyscale intensity, $\Delta I_1 = CI_1^a$ and $\Delta I_2 = CI_2^a$ denotes the minimum changes in perceived intensity at I_1 and I_2, respectively, referring to generalized Weber's model in Eq. (2.3).

Notice that the above Theorem is a generalization of the invariance result of Theorem 3.2.1. Consequently, for any $0 < a < 1$, the generalized Weber-based L^1 metrics constructed by density functions $\rho_a(y) = 1/y^a$ between the functions u and v, are given by,

$$D_a(u, v; \nu_a) = \int_X |u(x)^{-a+1} - v(x)^{-a+1}| dx. \tag{4.40}$$

Previously, we also presented a sketch of the proof of Theorem 4.2.1. Now, in order to determine the more detailed asymptotic behaviour of density functions that accommodate generalized Weber's model of perception, let us consider the forward problem: For any $a \in (0, 1)$, given any intensity $x = Cx^a \geq x_0$, where x_0 is a lower positive greyscale value, find a density function $g_a(y)$ such that the function (i.e., Eq. (4.28)),

$$G(x) = \int_x^{x+Cx^a} g_a(y) dy = K, \tag{4.41}$$

for all $x \geq x_0$, where K is a constant which represents the "visual accumulation" of intensity that contributes to human visual perception.

Again, from the assumption that $g_a(y)$ is a continuous function, we obtain the same equation as Eq. (4.29) by differentiating both sides of Eq. (4.41) with respect to x,

$$g_a(x + Cx^a)(1 + aCx^{a-1}) - g_a(x) = 0. \tag{4.42}$$

Note that Eq.(4.42) is a functional equation for the unknown $g(x)$, and it is unclear how to solve the equation explicitly. However, as previously introduced in section 3.2.1, we tried performing some asymptotic analysis on Eq. (4.42) by assuming a solution of the following form,

$$g_a(y) \sim y^\alpha, \tag{4.43}$$

where α is also a constant that is related to the constant a. By employing Eq. (4.43) on Eq. (4.42), we have

$$(x + Cx^a)^\alpha(1 + aCx^{a-1}) - x^\alpha = 0. \tag{4.44}$$

Case 1: The asymptotic limit of density functions as $x \to \infty$

For any $0 < a < 1$, we are now analyzing the asymptotic limit of $g_a(y)$ for $x \to \infty$. As $x \to \infty$, x dominates x^a in the first parenthesis of Eq. (4.44), thus we factor the equation as follows,

$$x^\alpha(1 + Cx^{a-1})^\alpha(1 + aCx^{a-1}) - x^\alpha = 0. \tag{4.45}$$

Since $0 < a < 1$, $x^{a-1} \to 0$ as $x \to \infty$. As such, we employ the binomial theorem expansion on the term $(1 + Cx^{a-1})^\alpha$ to obtain,

$$x^\alpha(1 + \alpha Cx^{a-1} + \cdots)(1 + aCx^{a-1}) - x^\alpha = 0, \tag{4.46}$$

which yields

$$(\alpha + a)Cx^{\alpha+a-1} + \alpha aC^2x^{\alpha+2(a-1)} + \cdots = 0, \tag{4.47}$$

after some manipulation. In order to eliminate the first term which is dominant in Eq. (4.47) as $x \to \infty$, we let

$$\alpha = -a \implies g_a(y) \sim y^{-a}, \tag{4.48}$$

which completes the proof of Theorem 4.1.2.

In order to make density functions $g_a(y)$ fitted more accurately in the equal area condition in Definition 3.2.1, we may wonder how to derive the next-to-leading order terms of the density function $g_a(y)$. Intuitively, one may suggest to carry out the binomial approximation to more terms to remove the higher order terms in Eq. (4.47). Hence, let us

perform more asymptotic analysis on the density function by first substituting $g_a(y) = y^\alpha$ into Eq. (4.41) and yielding

$$\begin{aligned} G(x) &= \int_x^{x+Cx^a} y^\alpha dy \\ &= \frac{1}{\alpha+1} y^{\alpha+1}\Big|_x^{x+Cx^a} \\ &= \frac{1}{\alpha+1}\Big[(x+Cx^a)^{\alpha+1} - x^{\alpha+1}\Big]. \end{aligned} \tag{4.49}$$

Once again, x dominates x^a in the term $x + Cx^a$, so we factor Eq. (4.49) and obtain

$$\int_x^{x+Cx^a} y^\alpha dy = \frac{1}{\alpha+1}\Big[x^{\alpha+1}(1+Cx^{a-1})^{\alpha+1} - x^{\alpha+1}\Big]. \tag{4.50}$$

Note $0 < a < 1$, Cx^{a-1} is sufficiently small as $x \to \infty$. Therefore, by employing the binomial expansion on the term $(1+Cx^{a-1})^{\alpha+1}$, Eq. (4.50) becomes

$$\int_x^{x+Cx^a} y^\alpha dy = \frac{1}{\alpha+1}\left[x^{\alpha+1}\left(1 + (\alpha+1)Cx^{a-1} + \frac{\alpha(\alpha+1)}{2}C^2x^{2a-2} + \cdots\right) - x^{\alpha+1}\right]. \tag{4.51}$$

After some simplification, we have

$$\int_x^{x+Cx^a} y^\alpha dy = Cx^{a+\alpha} + \frac{\alpha}{2}C^2x^{2a+\alpha-1} + \cdots. \tag{4.52}$$

Clearly, the first term in Eq. (4.52), i.e., $Cx^{a+\alpha}$, is dominant as $x \to \infty$, since $2a-b-1 = (a-b)+(a-1) < a-b$. In other words, higher order terms of Eq. (4.52) is subdominant since they involve increasing power of $a - 1 < 0$. As such, we have the following result: For $x > 0$, $0 < a < 1$, $\alpha < 0$ and $\alpha \neq -1$,

$$\int_x^{x+Cx^a} y^\alpha dy \sim Cx^{a+\alpha} \quad \text{as} \quad x \to \infty. \tag{4.53}$$

Notice that the condition $\alpha \neq -1$ originally comes from the antiderivative of y^α in Eq. (4.49), and $\alpha < 0$ is due to the fact that the density function accommodating Weber's generalized model of perception is a decreasing function.

In the special case that $\alpha = -1$, Eq. (4.49) becomes

$$\begin{aligned} G(x) &= \int_x^{x+Cx^a} y^{-1} dy \\ &= \ln(x+Cx^a) - \ln(x) \\ &= \ln(1+Cx^{a-1}). \end{aligned} \tag{4.54}$$

As $x \to \infty$, $Cx^{a-1} \to 0$, since $0 < a < 1$. By Taylor series approximation, Eq. (4.54) can be expanded as follows,

$$\int_x^{x+Cx^a} y^{-1} dy = Cx^{a-1} - \frac{1}{2}C^2 x^{2a-2} + \cdots, \tag{4.55}$$

which coincides with Eq. (4.52) for $\alpha = -1$. Overall, we can set up the following important lemma.

Lemma 4.2.1: For any $x > 0$, $0 < a < 1$ and $\alpha < 0$,

$$\int_x^{x+Cx^a} y^{\alpha} dy \sim Cx^{a+\alpha} \quad \text{as} \quad x \to \infty. \tag{4.56}$$

Note that when $\alpha = -a$, Lemma 4.2.1 yields the result,

$$\int_x^{x+Cx^a} y^{-a} dy \sim C \quad \text{as} \quad x \to \infty, \tag{4.57}$$

which is in agreement with the conclusion in Theorem 4.1.1.

Besides, let us now proceed further and find the second and third leading order terms of the density functions $g_a(y)$, satisfying the invariance result in Eq. (4.41), for $0 < a < 1$ as $y \to \infty$. First of all, by substituting $\alpha = -a$ into Eq. (4.52), we obtain

$$\int_x^{x+Cx^a} y^{-a} dy - C = -\frac{a}{2}C^2 x^{a-1} + O(x^{2a-2}). \tag{4.58}$$

Up to a constant, the first term in the RHS of Eq. (4.58) matches up the leading order term in the asymptotic expansion of $G(x)$ in Eq. (4.55) in the case $\alpha = -1$. Multiplying by the appropriate constant i.e., $-\frac{1}{2}aC$, on Eq. (4.55) yields,

$$-\frac{1}{2}aC \int_x^{x+Cx^a} y^{-1} dy = -\frac{1}{2}aC^2 x^{a-1} + \frac{1}{4}aC^3 x^{2a-2} + O(x^{3a-3}). \tag{4.59}$$

The leading order behaviour of the RHS in Eq. (4.59) is consistent with the dominant behaviour of the RHS in (4.58). By subtracting Eq. (4.59) from Eq. (4.58), we obtain the equation as follows,

$$\int_x^{x+Cx^a} \left[\frac{1}{y^a} + \frac{1}{2}aC\frac{1}{y}\right] dy - C = O(x^{2a-2}) \quad \text{as} \quad x \to \infty. \tag{4.60}$$

As a result, the difference between the integral and the constant C is now $O(x^{2a-2})$ rather than $O(x^{a-1})$ in Eq. (4.58), which leads to the conclusion that $\frac{1}{y^a}+\frac{1}{2}aCy^{-1}$ is an improved approximation to the density function $g_a(y)$ in the limit $y \to \infty$, for any $0 < a < 1$. We can write that, up to an overall constant,

$$g_a(y) \sim \frac{1}{y^a} + \frac{1}{2}aC\frac{1}{y} \quad \text{as} \quad x \to \infty. \tag{4.61}$$

Indeed, one may realize that the crucial step to work out the second leading order behaviour of the density function $g_a(y)$ is to find the value of α, yielding the leading order term that corresponds to the dominant behaviour of the RHS of Eq. (4.58).

Following the same logic, we first specifically calculate the leading order behaviour of the RHS of Eq. (4.60) and rewrite it in the following,

$$\int_x^{x+Cx^a} \left[\frac{1}{y^a}+\frac{1}{2}aCy^{-1}\right]dy - C = \left[\frac{1}{6}a(a+1)-\frac{1}{4}a\right]C^3x^{2a-2}+O(x^{3a-3}) \quad \text{as} \quad x \to \infty. \tag{4.62}$$

To eliminate the first term in the RHS of Eq. (4.62), imposing Eq. (4.52) and letting $\alpha = a - 2$ yields,

$$\int_x^{x+Cx^a} y^{a-2}dy = Cx^{2a-2} + \frac{1}{2}(a-2)C^2x^{3a-3} + \cdots. \tag{4.63}$$

Note that we have $\alpha = a - 2$ because we set $a + \alpha = 2a - 2$ to match the dominant behaviours in the RHS of Eq. (4.52) and (4.62) with each other. Multiplication by the appropriate constant, namely $[\frac{1}{6}a(a+1) - \frac{1}{4}a]C^2$, on Eq. (4.63), yields,

$$\left[\frac{1}{6}a(a+1)-\frac{1}{4}a\right]C^2\int_x^{x+Cx^a} y^{a-2}dy = \left[\frac{1}{6}a(a+1)-\frac{1}{4}a\right]C^3x^{2a-2}+O(x^{3a-3})+\cdots. \tag{4.64}$$

By subtracting Eq. (4.64) from Eq. (4.62), we have that

$$\int_x^{x+Cx^a} \left[\frac{1}{y^a}+\frac{1}{2}aC\frac{1}{y}+\left[\frac{1}{4}a-\frac{1}{6}a(a+1)\right]C^2y^{a-2}\right]dy - C = O(x^{3a-3}) \quad \text{as} \quad x \to \infty. \tag{4.65}$$

As $x \to \infty$, $O(x^{3a-3})$ goes to zero faster than $O(x^{2a-2})$, which implies the following equation is the better approximation to the density function $g_a(y)$ meeting the invariance requirement in Eq. (4.41) according to Weber's generalized model of perception,

$$g_a(y) \sim \frac{1}{y^a} + \frac{1}{2}aC\frac{1}{y} + \left[\frac{1}{4}a-\frac{1}{6}a(a+1)\right]C^2y^{a-2} \quad \text{as} \quad x \to \infty. \tag{4.66}$$

In principle, we can continue the procedure above and find infinitely more terms for the asymptotic expansion of the density function $g_a(y)$. If we consider $g_a(y) = 1/y^a$ is the zeroth leading order term, the value of α used in Eq. (4.52) to find the n^{th} leading order behaviour is

$$\alpha = n(a-1) - a, \tag{4.67}$$

which leads to an net asymptotic expansion of $g_a(y)$ having the following form,

$$g_a(y) = \sum_{n=0}^{\infty} a_n x^{n(a-1)-a} \quad \text{as} \quad x \to \infty. \tag{4.68}$$

In summary, let us list the following steps [13] involved in the derivation of Eq. (4.68).

Step 1: By setting $b = a$ in Eq. (4.52), we obtained, after a little rearrangement, the asymptotic result in Eq. (4.58), which implies that the invariance result according to Eq. (4.41) is satisfied, to leading order, by $g_a(y) \sim 1/y^a$, in agreement with the dominant behaviour introduced in Theorem 4.1.1.

Step 2: Secondly, we impose the result for $\alpha = -1$ in Eq. (4.55) in an effort to match the leading term on its RHS with the dominant behaviour of the RHS of Eq. (4.58), namely, $O(x^{a-1})$. Then, we multiply both sides of Eq. (4.55) by the appropriate constant so that the coefficients of the dominant terms in $O(x^{a-1})$ match. The term in the square bracket in Eq. (4.60) is the improved approximation of $g_a(y)$ as $y \to \infty$.

Step 3: To match the higher order terms in Eq. (4.52), namely, $x^{n(a-1)}$, with the leading order term $x^{a+\alpha}$ in Eq. (4.52), we employ the relation, $\alpha = n(a-1) - a$ for $n \geq 0$, on finding more leading order behaviours of the density function $g_a(y)$, which derives the general form of the asymptotic expansion of $g_a(y)$ in Eq. (4.68).

To conclude, in this case, we not only derived the general form of the asymptotic behaviour of density functions $g_a(y)$, but provided a method of computing the asymptotic expansion of $g_a(y)$ for $y \to \infty$ to any number of terms for the case $0 < a < 1$, by the above steps i.e., from step 1 to step 3.

Case 2: The asymptotic limit of density functions as $x \to 0^+$

In this case, we are looking for the asymptotic behaviour of $g_a(y)$ for any $0 < a < 1$ as $x \to 0^+$. Since $x \to 0^+$, x^a dominates x in the first parenthesis of Eq. (4.44), and we factor as follows,

$$[Cx^a(1 + C^{-1}x^{1-a})]^{\alpha}[aCx^{a-1}(1 + (aC)^{-1}x^{1-a})] - x^{\alpha} = 0. \tag{4.69}$$

After a little manipulation, we have,

$$aC^{\alpha+1}x^{a\alpha+a-1}[1+C^{-1}x^{1-a}]^{\alpha}[1+(aC)^{-1}x^{1-a}]-x^{\alpha}=0. \tag{4.70}$$

Note that the term $x^{1-a}\to 0$ as $x\to 0^{+}$ for any $0<a<1$. As such, we obtain the following equation by employing binomial expansion on the term $(1+C^{-1}x^{1-a})^{\alpha}$,

$$aC^{\alpha+1}x^{a\alpha+a-1}[1+\alpha C^{-1}x^{1-a}+\cdots][1+(aC)^{-1}x^{1-a}]-x^{\alpha}=0. \tag{4.71}$$

After some rearrangement, Eq. (4.71) becomes

$$aC^{\alpha+1}x^{a\alpha+a-1}\left[1+\frac{1}{C}(\alpha+\frac{1}{a})x^{1-a}+\cdots\right]-x^{\alpha}=0. \tag{4.72}$$

To eliminate the term x^{α}, we have to match it with the leading term $x^{a\alpha+a-1}$, which implies the following relation,

$$a\alpha+a-1=\alpha \quad\Longrightarrow\quad \alpha(a-1)+(a-1)=0. \tag{4.73}$$

Since $0<a<1$, we obtain $\alpha=-1$, which suggests that

$$g_a(y)\sim y^{-1}. \tag{4.74}$$

When $\alpha=-1$, $C^{\alpha+1}=C^{0}=1$. Thus, the constant C from the leading term on the left of Eq. (4.72) has been eliminated, but we still have the following,

$$(a-1)x^{-1}=0, \tag{4.75}$$

which is not satisfied with any $0<a<1$. The failure of this method suggests the assumption that $g_a(y)\sim y^{\alpha}$ as $y\to 0^{+}$ is invalid. To show the invalidity of the form, $g_a(y)\sim y^{\alpha}$ as $x\to 0^{+}$, more explicitly, let us first consider the following helpful lemmas.

Lemma 4.2.2: For any $x>0$, $0<a<1$ and $-1<\alpha<0$,

$$\int_{x}^{x+Cx^{a}} y^{\alpha}dy\sim\frac{C^{\alpha+1}}{\alpha+1}x^{a(\alpha+1)}\quad\text{as}\quad x\to 0^{+}. \tag{4.76}$$

Proof: First of all, we consider the case $\alpha\neq -1$. By basic calculation, we obtain,

$$\begin{aligned} G(x)&=\int_{x}^{x+Cx^{a}} y^{\alpha}dy\\ &=\frac{1}{\alpha+1}y^{\alpha+1}\Big|_{x}^{x+Cx^{a}}\\ &=\frac{1}{\alpha+1}\left[(x+Cx^{a})^{\alpha+1}-x^{\alpha+1}\right]. \end{aligned} \tag{4.77}$$

Because $x \to 0^+$, x^a dominates x, then we rewrite Eq. (4.77) as follows,

$$G(x) = \frac{1}{\alpha+1}\left[(Cx^a)^{\alpha+1}\left(1+\frac{1}{C}x^{1-a}\right)^{\alpha+1} - x^{\alpha+1}\right]. \tag{4.78}$$

Since $0 < a < 1$ and x^{1-a}/C is sufficiently small as $x \to 0^+$, employing the binomial approximation on the term $(1+\frac{1}{C}x^{1-a})^{\alpha+1}$ yields,

$$G(x) = \frac{1}{\alpha+1}\left[(Cx^a)^{\alpha+1}\left(1+(\alpha+1)\frac{1}{C(1+\alpha)}x^{1-a}+\frac{1}{2C}\alpha(1+\alpha)x^{2(1-a)}+\cdots\right) - x^{\alpha+1}\right]. \tag{4.79}$$

For $0 < a < 1$, it is clear that $0 < a(1+\alpha) < (1+\alpha)$, which determines the dominant behaviour of the RHS as $x \to 0^+$ and completes the proof.

Lemma 4.2.3: For any $x > 0$, $0 < a < 1$ and $\alpha = -1$,

$$\int_x^{x+Cx^a} y^{-1}dy \sim (a-1)\ln x \quad \text{as} \quad x \to 0^+. \tag{4.80}$$

Proof: From Eq. (4.54), we have

$$\begin{aligned} G(x) &= \int_x^{x+Cx^a} y^{-1}dy \\ &= \ln(1+Cx^{a-1}) \\ &= \ln(Cx^{a-1}) + \ln\left(1+\frac{1}{C}x^{1-a}\right) \\ &= (a-1)\ln x + \ln C + \frac{1}{C}x^{1-a} - \frac{1}{2C}x^{2(1-a)} + \cdots. \end{aligned} \tag{4.81}$$

where we have used the Taylor approximation of $\ln\left(1+\frac{1}{C}x^{1-a}\right)$. Obviously, the leading order term of $G(x)$ is $(a-1)\ln x$ for $\alpha = -1$. The proof is complete.

Lemma 4.2.4: For any $x > 0$, $0 < a < 1$ and $\alpha < -1$,

$$\int_x^{x+Cx^a} y^{\alpha}dy \sim -\frac{1}{\alpha+1}x^{\alpha+1} \quad \text{as} \quad x \to 0^+. \tag{4.82}$$

Proof: With reference to Eq. (4.79) in the proof of Lemma 4.1.2, it follows that $1+\alpha < a(1+\alpha) < 0$ in that $0 < a < 1$ and $\alpha < -1$, which implies the term $x^{\alpha+1}$ is dominant as $x \to 0^+$.

From the aforementioned three lemmas, we obtain the following results,

- If $-1 < \alpha < 0$, then

$$\int_x^{x+Cx^a} y^\alpha dy \;\to\; 0 \quad \text{as} \quad x \to 0^+. \tag{4.83}$$

- If $\alpha < -1$, then

$$\int_x^{x+Cx^a} y^\alpha dy \;\to\; \infty \quad \text{as} \quad x \to 0^+. \tag{4.84}$$

Essentially, our goal is to find the value of α such that the equal area condition in Eq. (4.41) holds as $x \to 0^+$. Whereas, we now notice no value of $\alpha \in (-\infty, 0)$ works by the results in Eq. (4.83) and (4.84). This implies that y^α is not the correct leading order behaviour of the density function $g_a(y)$ for $x \to 0^+$. Naturally, the next step is to investigate on the appropriate form of leading order term of $g_a(y)$, satisfying the invariance results in Eq. (4.41), as $x \to 0^+$. Before that, let us first illustrate the following significant theorem [13] that describes the relation between the equal area condition on generalized Weber's model of perception and the associated density function $g_a(y)$ for $a > 0$.

Theorem 4.2.1: For given values of $a > 0$, $C > 0$ and $A > 0$, let $g_a(y)$ satisfy the invariance condition in Eq. (4.41). Then, the indefinite integral,

$$\int_A^\infty g_a(y)dy, \tag{4.85}$$

must diverge.

Proof: We let $y_0 = A$, and define the following sequence,

$$y_{n+1} = y_n + Cy_n^a, \tag{4.86}$$

for $n \geq 0$. From the invariance result in Eq. (4.41), it follows that for $n \geq 1$,

$$\int_{x_n}^{x_{n+1}} g_a(y)dy = \int_{x_0}^{x_1} g_a(y)dy = K, \tag{4.87}$$

where $K > 0$ is a constant. From Eq. (4.86), $y_1 = y_0 + Cy_0^a > y_0$, since a, C and y_0 are all positive. This implies that

$$y_2 = y_1 + Cy_1^a = y_0 + Cy_0^a + Cy_1^a > y_0 + 2Cy_0^a, \tag{4.88}$$

which suggests that

$$y_n > y_0 + nCy_0^a, \tag{4.89}$$

for $n \geq 1$. In fact, the claim in Eq. (4.89) can be proved by inductive arguments. As such, $y_n \to \infty$ as $n \to \infty$. Switching the index n in Eq. (4.87) to k and summing all the k from 0 to $n-1$ yields the following expression,

$$\sum_{k=0}^{n-1} \int_{x_k}^{x_{k+1}} g_a(y)dy = \sum_{k=0}^{n-1} \int_{x_0}^{x_1} g_a(y)dy = \sum_{k=0}^{n-1} K = nK, \tag{4.90}$$

which implies

$$\int_{x_0}^{x_n} g_a(y)dy = nK, \tag{4.91}$$

As $n \to \infty$, Eq. (4.90) becomes

$$\int_{x_0}^{\infty} g_a(y)dy = \infty, \tag{4.92}$$

which completes the proof.

Now, let us proceed and find the correct form of leading order behaviour of $g_a(y)$ referring to Eq. (4.41) for $0 < a < 1$ as $x \to 0^+$. If we denote the antiderivative of leading order term of $g_a(y)$ as $A(y)$, Eq. (4.41) can be rewritten as follows,

$$G(x) = \int_{x}^{x+Cx^a} g_d(y)dy = A(x + Cx^a) - A(x) \sim K, \tag{4.93}$$

where $g_d(y)$ and K represents the valid dominant behaviour of $g_a(y)$ and a positive constant, respectively. For any $0 < a < 1$ and $C > 0$, both x and $x + Cx^a$ approach 0 as $x \to 0^+$, which implies $A(x) \to A(0)$ and $A(x + Cx^a) \to A(0)$. Interestingly, the difference between $A(x + Cx^a)$ and $A(x)$ is approaching to a positive constant rather than 0 or ∞, according to Eq. (4.93). This might make one consider leading order terms which diverge as $y \to 0^+$ but which involve logarithms. In what follows, we are going to test a few functions and see whether Eq. (4.93) is satisfied in terms of the functions for $0 < a < 1$ as $x \to 0^+$.

- $g_d(y) = -\ln y > 0$ for $0 < y < 1$ so that $g(y) \to \infty$ as $y \to 0^+$. In this case, $A(y) = y - y\ln y$, since

  $$\int \ln y \; dy = y \ln y - y. \tag{4.94}$$

  As a result, we have

  $$A(x + Cx^a) - A(x) = Cx^a - (x + Cx^a)\ln(x + Cx^a) + x\ln(x) \to 0, \tag{4.95}$$

  as $x \to 0^+$.

- $g_d(y) = -\dfrac{1}{y \ln y} > 0$ for $0 < y < 1$ so that $g(y) \to \infty$ as $y \to 0^+$. It is clear that $A(y) = -\ln|\ln y|$, which implies that

$$A(x + Cx^a) - A(x) = -\ln|\ln(x + Cx^a)| + \ln|\ln x|. \tag{4.96}$$

From Eq. (4.81), we obtain

$$\begin{aligned} A(x + Cx^a) - A(x) &= -\ln\left| a \ln x + \ln C + \frac{1}{C} x^{1-a} + O(x^{2-2a}) \right| + \ln|\ln x| \\ &= -\ln\left| -a + \frac{\ln C}{|\ln x|} + \frac{1}{C}\frac{x^{1-a}}{|\ln x|} + \frac{O(x^{2-2a})}{|\ln x|} \right| \\ &\to \ln\left(\frac{1}{a}\right), \end{aligned} \tag{4.97}$$

as $x \to 0^+$ for any $a, x, x + Cx^a \in (0, 1)$.

Note $\ln(1/a)$ is a positive constant. Indeed, Eq. (4.97) suggests that $g_d(y) = -\dfrac{1}{y \ln y}$ is a valid leading order term which satisfies Eq. (4.93) as $y \to 0^+$. Up to some constant, the leading order behaviour of the density function $g_a(y)$ as $y \to 0^+$ can be stated as follows,

$$g_a(y) \simeq -\frac{1}{y \ln y}. \tag{4.98}$$

Although, in practice, only finite intervals $[A, B] \subset \mathbb{R}_g$ of the intensity space are employed, as opposed to the infinite interval $[0, \infty)$; the asymptotic analysis on studying the leading order terms of the density function is still interesting as Eq. (4.98) presents the dominant behaviour of the density function, from which we can establish the generalized Weber-based measures in terms of very low intensities.

Finally, we skip the further asymptotic analysis of density functions as $x \to 0^+$ for $0 < a < 1$. One could follow similar principles, which are step 1, 2 and 3, as showed in the previous case. The Taylor approximation of the logarithm would be applied appropriately during the process.

4.2.2 The "Reverse Problem" for Weber's Generalized Model of Perception

Before investigating the asymptotic behaviour of the density functions that accommodate generalized Weber's model of perception for the other case $a > 1$, let us first consider

the "reverse problem" to verify previous asymptotic results of the density functions and provide a few connections to the other asymptotic analysis which will be introduced in the next section. In "reverse problem", given the density function $g(y) = y^{-a}$, we aim to find the leading order behaviour of $f(x)$ such that

$$F(x) = \int_x^{x+Cf(x)} \frac{1}{y^a} dy = K > 0, \tag{4.99}$$

where K is a constant which again represents the "visual accumulation" of intensity that contributes to human visual perception; C, a positive constant, is usually called Weber's constant. By differentiating Eq. (4.97) with respect to x, we obtain the following differential equation (DE) for $f(x)$ after a little manipulation,

$$1 + Cf'(x) = \left[\frac{x + Cf(x)}{x}\right]^a. \tag{4.100}$$

Indeed, Eq. (4.100) can be written as the DE,

$$\omega'(x) = \left[\frac{\omega(x)}{x}\right]^a, \tag{4.101}$$

where

$$\omega(x) = x + Cf(x). \tag{4.102}$$

When $a = 1$, we easily obtain $f(x) = kx = kx^1$, where k is a constant. Since the constant k can be merged into the Weber's constant C, we then simply have $f(x) = x = x^1$, which corresponds to standard Weber's model, i.e., $\Delta x/x = C$. For $a > 0$ and $a \neq 1$, solving the separable DE in Eq. (4.101) yields

$$\omega^{-a+1} = x^{-a+1} + A, \tag{4.103}$$

where A denotes an arbitrary constant. From Eq. (4.103), we have

$$\omega(x) = [x^{-a+1} + A]^{1/(1-a)}, \tag{4.104}$$

Substituting Eq. (4.104) into Eq. (4.102) yields

$$Cf(x) = \omega(x) - x = [x^{-a+1} + A]^{1/(1-a)} - x, \tag{4.105}$$

We now wish to determine the dominant behaviour of $f(x)$ as $x \to \infty$.

Case 1: For $0 < a < 1$, the term x^{-a+1} in Eq. (4.104) will dominate, so we factor as follows,

$$\omega(x) = x[1 + Ax^{a-1}]^{1/(1-a)}, \tag{4.106}$$

which implies that

$$Cf(x) = x[1 + Ax^{a-1}]^{1/(1-a)} - x. \tag{4.107}$$

Once again, as $x \to \infty$, $x^{a-1} \to 0$, which allows us to impose binomial approximation to expand the term in square brackets in Eq. (4.107) and obtain,

$$Cf(x) \simeq x\left[1 + \frac{A}{(1-a)}x^{a-1}\right] - x = \frac{A}{(1-a)}x^a. \tag{4.108}$$

In fact, Eq. (4.108) suggests that $f(x) \sim x^a$, which is in agreement with the results, i.e., $\Delta I_1 = CI_1^a$ and $\Delta I_2 = CI_2^a$, in Theorem 4.1.1.

Case 2: For $a > 1$, the term x^{-a+1} in the square brackets in Eq. (4.102) will approach zero as $x \to \infty$. As such, we factor out the A,

$$\omega(x) = A^{1/(1-a)}[1 + A^{-1}x^{a-1}]^{1/(1-a)}. \tag{4.109}$$

By employing the binomial approximation,

$$\begin{aligned} \omega(x) &\simeq A^{1/(1-a)}\left[1 + \frac{1}{A(1-a)}x^{-a+1}\right] \\ &= A^{1/(1-a)} + \frac{A^{1/(1-a)}}{A(1-a)}x^{-a+1}. \end{aligned} \tag{4.110}$$

As $a > 1$, we have that

$$\omega(x) \to A^{1/(1-a)} \quad \text{as} \quad x \to \infty. \tag{4.111}$$

Note that the function $\omega(x)$ is the upper limit of integration to the function $F(x)$ in Eq. (4.99), which implies that

$$F(x) = \int_x^{\omega(x)} \frac{1}{y^a}dy = K > 0. \tag{4.112}$$

For $a > 1$, we now find the constancy of the function $F(x)$ is satisfied, at least as $x \to \infty$, if, to leading order,

$$F(x) = \int_x^B \frac{1}{y^a}dy, \tag{4.113}$$

where $B = A^{1/(1-a)}$. Instead of $w(x)$ getting larger and larger as $x \to \infty$, $\omega(x)$ is decreasing towards the constant B, in which case $F(x)$ is negative. This actually, once again, implies that $1/y^a$ is not a correct form of the asymptotic behaviour of density functions. Moreover, the incorrectness of the result that $g_a(y) \sim 1/y^a$ as $y \to \infty$, for $a > 1$, can be explicated by Theorem 4.2.1 in that

$$F(x) = \int_x^B \frac{1}{y^a} dy < \infty. \tag{4.114}$$

4.2.3 Asymptotic analysis of density functions for the case $a > 1$

In Section 4.2.1, we have discussed the asymptotic expansion of density functions for $0 < a < 1$. It now remains to perform an asymptotic analysis of density functions for the case $a > 1$. As a matter of fact, we shall show that the asymptotic expansion of density functions for $a > 1$ as $x \to 0^+$ ($0 < a < 1$ as $x \to \infty$) employs the same expression as those used for the asymptotic limit $x \to \infty$ for the case $0 < a < 1$ ($x \to 0^+$ for the case $a > 1$).

Case 3: The asymptotic limit of density functions as $x \to \infty$

In this case, we present that for $a > 1$, the asymptotic behaviour of the density function $g_a(x)$, $x \to \infty$, cannot be represented by a power law. Similar to Lemmas 4.2.2, 4.2.3 and 4.2.4 that were previously introduced in Case 2 in Section 4.2.1, we show the following three important lemmas to analyze the asymptotic limit for this case.

Lemma 4.2.5: For any $x > 0$, $a > 1$ and $\alpha < -1$,

$$\int_x^{x+Cx^a} y^\alpha dy \sim -\frac{1}{\alpha+1} x^{\alpha+1} \quad \text{as} \quad x \to \infty. \tag{4.115}$$

Proof: By straightforward calculation, we obtain,

$$G(x) = \int_x^{x+Cx^a} y^\alpha dy = \frac{1}{\alpha+1}\left[(x+Cx^a)^{\alpha+1} - x^{\alpha+1}\right]. \tag{4.116}$$

As $x \to \infty$, the term x^a dominates x. As a result, we may proceed in the following,

$$G(x) = \frac{1}{\alpha+1}\left[(Cx^a)^{\alpha+1}\left(1 + \frac{1}{C}x^{1-a}\right)^{\alpha+1} - x^{\alpha+1}\right]. \tag{4.117}$$

Since $a > 1$ and x^{1-a}/C approaches zero as $x \to \infty$, employing the binomial approximation on the term $(1 + \frac{1}{C}x^{1-a})^{\alpha+1}$ yields,

$$G(x) = \frac{1}{\alpha+1}\left[(Cx^a)^{\alpha+1}\left(1 + (\alpha+1)\frac{1}{C(1+\alpha)}x^{1-a} + \frac{1}{2C}\alpha(1+\alpha)x^{2(1-a)} + \cdots\right) - x^{\alpha+1}\right]. \tag{4.118}$$

Note that the final line is identical to Eq. (4.79) in the proof of Lemma 4.2.2. Since $a > 1$, the exponent $a(1+\alpha) < 1+\alpha$, which implies that the term $x^{1+\alpha}$ dominates as $x \to \infty$, yielding the desired result.

Lemma 4.2.6: For any $x > 0$, $a > 1$ and $\alpha = -1$,

$$\int_x^{x+Cx^a} y^{-1}dy \sim (a-1)\ln x \quad \text{as} \quad x \to \infty. \tag{4.119}$$

Proof:

$$\begin{aligned}\int_x^{x+Cx^a} y^{-1}dy &= \ln(1 + Cx^{a-1}) \\ &= \ln(Cx^{a-1}) + \ln\left(1 + \frac{1}{C}x^{1-a}\right) \\ &= (a-1)\ln x + \ln C + \frac{1}{C}x^{1-a} - \frac{1}{2C}x^{2(1-a)} + \cdots.\end{aligned} \tag{4.120}$$

where we impose the Taylor series expansion of $\ln\left(1 + \frac{1}{C}x^{1-a}\right)$. The proof is complete.

Lemma 4.2.7: For any $x > 0$, $a > 1$ and $-1 < \alpha < 0$,

$$\int_x^{x+Cx^a} y^{\alpha}dy \sim \frac{C^{\alpha+1}}{\alpha+1}x^{a(\alpha+1)} \quad \text{as} \quad x \to \infty. \tag{4.121}$$

Proof: From Eq. (4.118) in the proof of Lemma 4.2.5, for $a > 1$ and $-1 < \alpha < 0$, it follows that $0 < 1 + \alpha < a(1+\alpha)$, which suggests that the term $x^{a(\alpha+1)}$ is dominant as $x \to \infty$. This completes the proof.

From Lemmas 4.2.5, 4.2.6 and 4.2.7, we can conclude the following for the case $a > 1$:

- If $-1 \leq \alpha < 0$, then

$$\int_x^{x+Cx^a} y^{\alpha}dy \;\to\; \infty \quad \text{as} \quad x \to \infty. \tag{4.122}$$

- If $\alpha < -1$, then

$$\int_x^{x+Cx^a} y^\alpha dy \;\to\; 0 \quad \text{as} \quad x \to \infty. \tag{4.123}$$

From Eqs. (4.122) and (4.123) with the assumption that the leading order behaviour of the density function $g_a(y)$ for $y \to \infty$ is of the form $g_a(y) \sim y^\alpha$, we found there is no α value for which the integral approaches a constant as $x \to \infty$. This leads us to thinking about the new form of leading order behaviour of density functions $g_a(y)$ for $y \to \infty$. Recall that this was also the case for the asymptotic limit of density functions in the case $0 < a < 1$ as $x \to 0^+$ – see Eqs. (4.83) and (4.84). Similarly, one may consider leading order terms which converge to zero as $y \to \infty$ but which involve logarithms. Based on our previous analysis in Case 2, we examine the following form of the density function,

$$g(y) = \frac{1}{y \ln y} > 0 \quad \text{for} \quad y > 1. \tag{4.124}$$

Note that $g(y) \to 0$ as $y \to \infty$. To verify, let us integrate $g(y)$ in Eq. (4.124) from x to $x + Cx^a$ and obtain,

$$\begin{aligned}
\int_x^{x+Cx^a} \frac{1}{y \ln y} dy &= \ln[\ln(x + Cx^a)] - \ln(\ln x) \\
&= \ln\left[a \ln x + \ln C + \frac{1}{C} x^{1-a} + O(x^{2-2a})\right] - \ln(\ln x) \\
&= \ln\left[a + \frac{\ln C}{\ln x} + \frac{1}{C}\frac{x^{1-a}}{\ln x} + \frac{O(x^{2-2a})}{\ln x}\right] \\
&\to \ln(a),
\end{aligned} \tag{4.125}$$

as $x \to \infty$. As such, we may view the integrand in Eq. (4.125) as an leading order behaviour of the density function $g_a(y)$ for $a > 1$, i.e.,

$$g_a(x) \simeq \frac{1}{x \ln x} \quad \text{as} \quad x \to \infty. \tag{4.126}$$

By comparing Eq. (4.126) with Eq. (4.98), we realize that, up to some constant, the leading order term of density functions $g_a(y)$ for the case $0 < a < 1$ as $x \to 0^+$ is the same as the dominant behaviour of $g_a(y)$ for $a > 1$ as $x \to \infty$.

Case 4: The asymptotic limit of density functions as $x \to 0^+$

In Case 3, we showed that an analysis of the asymptotic limit $x \to \infty$ for the case $a > 1$ employs the same equations as those used for the asymptotic limit $x \to 0^+$ for the case

$0 < a < 1$. As such, we expect that the asymptotic behaviour of density functions for the case $a > 1$ as $x \to 0^+$ corresponds to the identical form of the asymptotic limit presented in the case $0 < a < 1$ as $x \to \infty$.

Indeed, if we set $\alpha = -a$ and substitute it into Eq. (4.51), we arrive at Eq. (4.58) which may also be viewed as an asymptotic relation for $a > 1$ as $x \to 0^+$,

$$\int_x^{x+Cx^a} \frac{1}{y^a} dy - C = \frac{1}{2} aC^2 x^{a-1} + O(x^{2a-2}). \tag{4.127}$$

Then, we can employ Steps 1, 2 and 3 as introduced in Case 1 to produce an asymptotic expansion of density functions in the following,

$$g_a(x) = \sum_{n=0}^{\infty} \frac{a_n}{x^{a+n(1-a)}} \quad \text{as} \quad x \to 0^+, \tag{4.128}$$

which is exactly the same form as Eq. (4.68) showed in Case 1 for $0 < a < 1$ as $x \to \infty$. Note as n increases for $n \geq 1$, the term $x^{a+n(1-a)}$ subdominate the term x^a as $x \to 0^+$ since $a > 1$ and the exponents $n(1-a)$ are negative.

Finally, by combining all asymptotic bahaviours of density functions $g_a(y)$ in aforementioned cases, the significant results can be stated to give

$$\begin{aligned}
0 < a < 1: \quad & g_a(y) \simeq \frac{1}{y \ln y} \quad \text{as} \quad y \to 0^+. \\
a > 1: \quad & g_a(y) = \frac{1}{y^a} + \frac{1}{2} aC \frac{1}{y} - \frac{1}{12} aC^2 (2a-1) \frac{1}{y^{2-a}} + \cdots \text{as} \quad y \to 0^+. \\
0 < a < 1: \quad & g_a(y) = \frac{1}{y^a} + \frac{1}{2} aC \frac{1}{y} - \frac{1}{12} aC^2 (2a-1) \frac{1}{y^{2-a}} + \cdots \quad \text{as} \quad y \to \infty. \\
a > 1: \quad & g_a(y) \simeq \frac{1}{y \ln y} \quad \text{as} \quad y \to \infty.
\end{aligned} \tag{4.129}$$

As mentioned in Section 3.2.2, Weber's law breaks down at very low intensities, and our main interests concentrate on the high-intensity range, i.e., the asymptotic case $y \to \infty$. In many applications, e.g., medical image processing, the range of the intensity values produced by the image is typically very large. As such, one could be more interested in the last two asymptotic results in Eq. (4.129), and it follows Eq. (3.12) and Eq. (3.13) (the intensity-based measure) that the associated distance functions are as follows,

$$\begin{aligned}
0 < a < 1: \quad & D_a(u, v; \nu_a) = \int_X |u(x)^{-a+1} - v(x)^{-a+1}| dx. \\
a > 1: \quad & D_1(u, v; \nu_1) = \int_X |\ln(\ln u(x)) - \ln(\ln v(x))| \, dx.
\end{aligned} \tag{4.130}$$

Chapter 5

Best Approximations and Stationarity Conditions for Generalized Weber-based and SSIM-based Approximations

In Section 3.3 of this thesis, we briefly considered the problem of function approximation using "Weberized" L^2-based metrics, that is, L^2 distance functions which employed intensity-based density functions of the form $\rho_a(y)$ for $0 < a < 1$. In Eq. (3.57), the stationarity conditions satisfied by the best approximations (in terms of the expansion coefficients c_k associated with a fixed orthonormal basis) were given. We begin this chapter (Section 5.1) with a more detailed analysis of the (finite-dimensional) optimization problem associated with this best approximation problem. In Section 5.2, we investigate the stationarity conditions satisfied by best generalized Weber-based approximations from a more theoretical, i.e., functional analysis, perspective involving the Fréchet derivative. These results have inspired a similar analysis of the best structural similarity-based (SSIM-based) approximation problem [17] which will be described in Section 5.3.

5.1 The problem of best function approximation for generalized Weber metrics

Previously, we showed that if the intensity-based measure ν supported on the greyscale

range space $\mathbb{R}_g$ is defined by a continuous density function $\rho_a(y)$ for $a > 0$, then the distance between two functions u and v is given by

$$D(u, v; \nu) = \int_X |P(u(x)) - P(v(x))| \, dx, \tag{5.1}$$

where P is the antiderivative of ρ_a, i.e., $P'(y) = \rho_a(y)$. As mentioned before, it is more convenient to work with the L^2-based analogues of the above distance functions, which are given by,

$$D_2(u, v; \nu) = \left[\int_X [P(u(x)) - P(v(x))]^2 \, dx \right]^{1/2}. \tag{5.2}$$

As mentioned earlier, we shall consider u to be a fixed reference function and v an approximation to it. We consider finite-dimensional approximations of the following standard form,

$$v_n(x) = \sum_{k=1}^{n} c_k \phi_k(x) \, , \tag{5.3}$$

where the set of functions $\{\phi_k\}_{k=1}^{\infty}$ is assumed to form a linearly independent or orthogonal/orthonormal basis on X. There is one complication, however, in that we must guarantee that the function $v_n(x)$ defined in Eq. (5.3) lies in the space of measurable and positive-valued functions $\mathcal{F}(X)$ which, we recall, is defined as follows,

$$\mathcal{F}(X) = \{u : X \to \mathbb{R}_g \mid 0 < A \leq u(x) \leq B \ \text{ a.e. } \ x \in X\}. \tag{5.4}$$

We define $C_n \in \mathbb{R}^n$ to be the feasible set of parameters, $\mathbf{c} = (c_1, \cdots, c_n)$ so that $v_n \in \mathcal{F}(X)$. In Section 6.2, we show that C_n is compact and convex. For now, we define the following subset of $S_n \subset \mathcal{F}(X)$,

$$S_n = \left\{ v : X \to \mathbb{R}_g \;\middle|\; v(x) = \sum_{k=1}^{n} c_k \phi_k(x) \ \text{ for } \ \mathbf{c} \in C_n \right\}. \tag{5.5}$$

In other words, $S_n = \mathcal{F}(X) \ \cap \ \text{span}\{\phi_k\}_{k=1}^{n}$.

In the problem of best approximation, our goal is to determine the optimal coefficients c_k in Eq. (5.3) that minimize the distance function $D_2(u, v_n; \nu)$ in Eq. (5.2). Let us now consider the squared D_2 distance function between u and v_n as a function of the parameters $c_1, \cdots, c_n$ and, for notational convenience, denote it as follows,

$$\begin{aligned} \Delta(\mathbf{c}) = \Delta(c_1, \cdots, c_n) &= \int_X [P(u(x)) - P(v_n(x))]^2 \, dx \\ &= \int_X [P(u(x)) - P(\sum_{k=1}^{n} c_k \phi_k(x))]^2 \, dx \, . \end{aligned} \tag{5.6}$$

We wish to find a set of optimal parameters $\mathbf{a} = (a_1, \cdots, a_n) \in C_n$ which minimizes $\Delta(\mathbf{c})$, i.e.,

$$\mathbf{a} = (a_1, \cdots, a_n) = \arg \min_{\mathbf{c} \in C_n} \Delta(\mathbf{c}). \tag{5.7}$$

If we assume that such a minimizer in Eq. (5.5) exists, then the function,

$$v_n = \sum_{k=1}^{n} a_k \phi_k(x) \in S_n \,, \tag{5.8}$$

will be considered to be the **best approximation** of u in S_n. It should be mentioned that the uniqueness of the minimizer may not be guaranteed since there might exist several local optimal solutions that minimize the distance function in Eq. (5.6). More details about the existence and uniqueness of the best approximation v_n will be discussed in the next chapter.

Now, let us consider the following finite-dimensional constrained optimization problem:

$$\min_{\mathbf{c} \in C_n} \Delta(\mathbf{c}), \tag{5.9}$$

where $\Delta(\mathbf{c})$, i.e., Eq. (5.6), is a differentiable function, and $C_n \subset \mathbb{R}^n$ is the feasible set that restricts all possible approximate functions v_n (i.e., Eq. (5.3)) in $\mathcal{F}(X)$. In what follows, we shall focus on necessary optimality conditions for the above problem (i.e., Eq. (5.9)). As an introduction, let us first review the definitions of the cone of internal directions [43], which are commonly dealt with specifying the optimality conditions in a number of optimization models.

Definition 5.1.1: Let $d \in \mathbb{R}^n$ and $c_0 \in C_n$. Then d is called the **internal direction** of C_n if and only if $\exists\ \alpha > 0$ and $\forall\ t \in [0, \alpha]$, we have

$$c_0 + td \in C_n. \tag{5.10}$$

In fact, the set of all internal directions of C_n at c_0 is a cone, which is denoted as $I(C_n, c_0)$.

Necessary Optimality Condition

Suppose that c_0 is a local solution to Eq. (5.9), then we have [44]

$$D\Delta(c_0)d \geq 0,\ \forall\ d \in I(C_n, c_0), \tag{5.11}$$

where $D\Delta(c_0)d$ denotes the directional derivative of the function $\Delta(\mathbf{c})$ at c_0 with the direction d.

Proof: From the definition of the internal cone, for any $d \in I(C_n, c_0)$, we have $c_0 + td \in C_n$ for all $t \in [0, \alpha]$, where $\alpha > 0$. As such, we may define a new function $g(t)$ on the interval $[0, \alpha]$ as follows,

$$g(t) = \Delta(c_0 + td). \tag{5.12}$$

Since c_0 is a minimizer of Δ, it is clear that 0 is a minimizer of g. Indeed, the first-order Taylor expansion for g at 0 is given by,

$$g(t) = g(0) + g'(0)t + o(t), \tag{5.13}$$

where $o(t)$ represents the higher-order terms of t which go to 0 faster than t as t approaches to 0. Now, we claim that

$$g'(0) \geq 0. \tag{5.14}$$

To show Eq. (5.14), suppose $g'(0) < 0$. Then there exists an $0 < \epsilon < \alpha$ such that for any $0 \leq t < \epsilon$, we have $o(t) < |g'(0)t|$, which implies that

$$g(t) < g(0) + g'(0)t + |g'(0)t|. \tag{5.15}$$

Since $g'(0) < 0$ and $0 \leq t < \epsilon$, the above equation can be rewritten as

$$g(t) < g(0) + g'(0)t - g'(0)t = g(0), \tag{5.16}$$

which contradicts that 0 is a minimum point of g. This implies that $g'(0) \geq 0$. From Eq. (5.12), we also have

$$g'(t) = D\Delta(c_0 + td)d. \tag{5.17}$$

Finally, setting $t = 0$ and combining the result with the claim, i.e., Eq. (5.14), yields that

$$D\Delta(c_0)d \geq 0, \tag{5.18}$$

which completes the proof.

5.2 Another look at the stationarity equations for best Weber-based approximation in terms of the Fréchet derivative

Let us now return to the optimization problem in Eq. (5.9), where $\Delta(\mathbf{c})$ is defined in Eq. (5.6). If we now impose the usual stationarity conditions on Eq. (5.6), we obtain

$$\frac{\partial \Delta}{\partial c_p} = -2 \int_X [P(u(x)) - P(v_n(x))] P'(v_n(x)) \phi_p(x) dx = 0, \quad 1 \leq p \leq n. \tag{5.19}$$

It then follows from $P'(y) = \rho_a(y)$ that

$$\int_X [P(u(x)) - P(v_n(x))] \rho_a(v_n(x)) \phi_p(x) dx = 0, \quad 1 \leq p \leq n. \tag{5.20}$$

In the standard Weber model, i.e., $P(y) = \log(y)$ for all $y \in \mathcal{F}(X)$, Eq. (5.20) will yield a complicated nonlinear system of equations in the unknowns c_k because of the nonlinearity of the logarithms. These complications will also exist for the Weber generalized models of perception as well. As introduced in Chapter 3, one may employ a gradient-descent scheme in deriving numerical estimates of the optimal coefficients a_k. For the initial guess of c_k in numerical experiments, it is reasonable to start with the best L^2 coefficients associated with the $\{\phi_k\}_{k=1}^n$ basis.

Let us now consider the stationarity equations in Eq. (5.20). For each $p \in \{1, 2, \cdots, n\}$, multiply both sides of Eq. (5.20) by a_p, the coefficient of the best approximation $v_n(x)$ to $u(x)$ in Eq. (5.8), and then sum over all p to obtain the following result,

$$\int_X [P(u(x)) - P(v_n(x))] \rho_a(v_n(x)) v_n(x) dx = 0. \tag{5.21}$$

However, the above operation is not restricted to the use of the optimal coefficients a_k. If we consider any set of coefficients $\{c_1, \cdots, c_n\}$ which define a function $h(x)$ of the form,

$$h(x) = \sum_{k=1}^{n} c_k \phi_k(x), \tag{5.22}$$

subject to the condition that $v_n + h \in \mathcal{F}(X)$, we arrive at the following equation,

$$\int_X [P(u(x)) - P(v_n(x))] \rho_a(v_n(x)) h(x) dx = 0. \tag{5.23}$$

Now, let us define the following vector space of functions,

$$H_n = \text{span}\{\phi_1, \cdots, \phi_n\}. \tag{5.24}$$

H_n is an n-dimensional subspace of $\mathbb{R}^n$. From Eq. (5.5), it is evident that $S_n \subset H_n$. Generally, $H_n \nsubseteq \mathcal{F}(X)$ since it will contain negative-valued functions. In addition, $\mathcal{F}(X) \nsubseteq H_n$ because $\mathcal{F}(X)$ is not even a vector space. For instance, if we have a function $u(x) \in \mathcal{F}(X)$. Then from Eq. (5.4), $2u(x) \in [2A, 2B]$ for a.e. $x \in X$, which implies that $2u(x) \notin \mathcal{F}(X)$. Therefore, $\mathcal{F}(X)$ is not a vector space. Nevertheless, we can show the function space $\mathcal{F}(X)$ is closed and bounded as follows.

Given any function $f \in \mathcal{F}(X) \subset L^2(X)$, we let

$$||f||_{a,2}^2 = \int_X [P(f(x))]^2 \, dx, \quad a \geq 0, \tag{5.25}$$

where $||\cdot||_{a,2}$ denotes the usual squared L^2 norm (i.e., $a = 0$ and $P(y) = y$) and modified squared L^2 metrics (i.e., generalized Weberized) , and P, as introduced before, is related to the intensity-based measures ν (i.e., Eq. (3.13)) supported on the greyscale range space $\mathbb{R}_g$. Since P is an increasing function on $\mathcal{F}(X)$, it then follows from Eqs. (3.13) and (5.4) that

$$P^2(A)m_g(X) \leq ||f||_{a,2}^2 \leq P^2(B)m_g(X), \quad a \geq 0, \tag{5.26}$$

where $m_g(X)$ denotes the Lebesgue measure of X. In particular, if $a = 0$, i.e., Lebesgue measure, and $X = [a, b]$, we have $\rho_0(y) = 1$ and $m_g(X) = b - a$. Clearly, Eq. (5.26) suggests that $\mathcal{F}(X)$ is a bounded set. For the closedness of $\mathcal{F}(X)$, let us take an arbitrary convergent sequence of functions $\{f_j\} \subset \mathcal{F}(X)$ such that

$$\lim_{j \to \infty} f_j = f. \tag{5.27}$$

Since $\{f_j\} \subset \mathcal{F}(X)$, we have that

$$A \leq f_j(x) \leq B, \quad \text{a.e. } x \in X, \tag{5.28}$$

for all $j \in \{1, 2, \cdots\}$, and the sequence of bounded functions uniformly converges to f, it then follows from Eqs. (5.27) and (5.28) that

$$A \leq f(x) \leq B, \quad \text{a.e. } x \in X. \tag{5.29}$$

Therefore, this suggests that $f \in \mathcal{F}(X)$, implying that $\mathcal{F}(X) \subset L^2(X)$ is a closed set.

Furthermore, it follows from Eqs. (5.22) and (5.23) that if $v_n \in S_n$ satisfies the stationarity conditions in Eq. (5.21), then Eq. (5.23) holds for all $h \in H_n$ such that $v_n + h \in S_n$.

Indeed, Eq. (5.23) has the appearance of a result involving Fréchet derivatives since the integral on the left side may be viewed as a linear operator acting on $h \in H_n$. Motivated by this result, we now consider the squared distance function in Eq. (5.6) in terms of the real-valued function $v \in \mathcal{F}$ and not the expansion coefficients c_k, i.e.,

$$\Delta(v) = \int_X [P(u(x)) - P(v(x))]^2 \, dx. \tag{5.30}$$

As such, we now consider $\Delta(v)$ to be a functional, i.e., $\Delta : \mathcal{F} \to \mathbb{R}$. From Eq. (5.23), we conjecture that the Fréchet derivative of $\Delta(v)$ is given by the following linear operator,

$$\Delta'(v)h = -2 \int_X [P(u(x)) - P(v(x))]\rho_a(v(x))h(x)dx. \tag{5.31}$$

From the necessary optimality condition mentioned in Section 5.1.1 , if $v_n \in S_n$ is a local minimizer of $\Delta(v)$, then

$$\Delta'(v_n)h \geq 0 \tag{5.32}$$

for all $h \in I(S_n, v_n) \subset H_n$, where $I(S_n, v_n)$ represents the cone of internal directions of S_n at v_n.

There are some complications, however, with the idea of Fréchet derivative since, technically speaking, our space $\mathcal{F}(X)$ is a metric space but not a normed space. That being said, we may be able to extend all of the usual definitions of Fréchet derivatives over normed spaces to our metric space $\mathcal{F}(X)$ which, in fact, has a linear structure, as far as linear combinations of functions are concerned. The fact that we are working with L^2-type distance functions also helps.

Special case: Lebesgue measure

To show the effectiveness of Eq. (5.31), let us now consider a special case of the D_2 metric defined in Eq. (5.2) in which $\rho(y) = 1$ and $P(y) = y$, i.e., Lebesgue measure. In this case, $\mathcal{F}(X)$ is a subset of $L^2(X)$ which is a complete set of functions. Once again, we consider u to be a reference function and v_n (i.e., 5.3) to be an approximation to it ,and work with the squared L^2 distance function Δ, defined as follows,

$$\Delta(v_n) = \|u - v_n\|_2^2 = \int_X [u(x) - v_n(x)]^2 \, dx. \tag{5.33}$$

For a function $h \in H_n \subset L^2(X)$ such that $v_n + h \in S_n$, the difference between $\Delta(v_n + h)$ and $\Delta(v_n)$ is given by

$$\Delta(v_n + h) - \Delta(v_n) = \int_X [u(x) - (v_n(x) + h(x))]^2 dx - \int_X [u(x) - v_n(x)]^2 \, dx, \tag{5.34}$$

which implies

$$\Delta(v_n + h) - \Delta(v_n) = -2 \int_X [u(x) - v_n(x)]h(x)dx + \int_X h(x)^2 \, dx. \tag{5.35}$$

Then, it follows from Eq. (5.31) that the Fréchet derivative of $\Delta(v_n)$ is stated as follows,

$$D\Delta h = \Delta'(v_n)h = -2 \int_X [u(x) - v_n(x)]h(x)dx, \tag{5.36}$$

which is linear in terms of h. From Eqs. (5.35) and (5.36), we have

$$\lim_{h \to 0} \frac{|\Delta(v_n + h) - \Delta(v_n) - D\Delta(v_n)h|}{\|h\|_2} = \lim_{h \to 0} \|h\|_2 = 0, \tag{5.37}$$

which conforms the valid form of the Fréchet derivative of $\Delta(v_n)$ derived in Eq. (5.36). As a result, if $\bar{v}_n$ is a stationary point of $\Delta(v_n)$, i.e.,

$$\Delta'(\bar{v}_n)h = -2 \int_X [u(x) - \bar{v}_n(x)]h(x)dx = 0 \quad \forall h \in H_n, \tag{5.38}$$

then from Eq. (5.35), for all $h \in H_n$, $h \neq 0$,

$$\Delta(\bar{v}_n + h) > \Delta(\bar{v}_n), \tag{5.39}$$

which suggests $\bar{v}_n$ is a minimizer of $\Delta(v_n)$.

In what follows, we shall try performing the same type of analysis on the more general D_2 metrics (i.e., Eq. (5.2)) defined by measures ν which correspond to density functions $\rho(y)$. As indicated in Chapter 4, we assume that $\rho(y)$ is a continuous function on $\mathbb{R}_g$. Given a reference function $u \in \mathcal{F}(X)$ and an approximation $v_n \in S_n$, for any $h \in H_n$ such that $v_n + h \in S_n$, according to Eq. (5.6), we examine the following difference,

$$\Delta(v_n + h) - \Delta(v_n), \tag{5.40}$$

as $h \to 0$. Note that, for convenience, we omit the subscript index of the approximate function v_n, and simply express it as v. First of all, it follows from Eq. (5.30) that

$$\Delta(v + h) = \int_X P(u(x))^2 dx - 2 \int_X P(u(x))P(v(x) + h(x))dx + \int_X [P(v(x) + h(x))]^2 dx. \tag{5.41}$$

Recall that $P'(y) = \rho(y)$, i.e., $P(y)$ is differentiable on $\mathbb{R}_g$. Then for a.e. $x \in X$, by the Mean Value Theorem, i.e., Theorem 3.2.2, we have,

$$\begin{aligned} P(v(x) + h(x)) &= P(v(x)) + P'(\xi(x))h(x) \\ &= P(v(x)) + \rho(\xi(x))h(x), \end{aligned} \tag{5.42}$$

where $\xi(x)$ lies between $v(x)$ and $v(x) + h(x)$. This implies that the second integral on the RHS of (5.41) may be written as follows,

$$\int_X P(u(x))P(v(x)+h(x))dx = \int_X P(u(x))P(v(x))dx + \int_X P(u(x))\rho(\xi(x))h(x)dx. \tag{5.43}$$

Similarly, the third integral on the RHS of Eq. (5.41) becomes

$$\int_X [P(v(x)+h(x))]^2 dx = \int_X P(v(x))^2 dx + 2\int_X P(v(x))\rho(\xi(x))h(x)dx + \int_X \rho(\xi(x))^2 h(x)^2 dx. \tag{5.44}$$

As such, the distance between $\Delta(v + h)$ and $\Delta(v)$ is as follows,

$$\Delta(v+h) - \Delta(v) = -2\int_X [P(u(x)) - P(v(x))]\rho(\xi(x))h(x)dx + \int_X \rho(\xi(x))^2 h(x)^2 dx. \tag{5.45}$$

Note that in the special case $\rho(y) = 1$, i.e., Lebesgue measure, the above equation becomes Eq. (5.35), as expected. The first integral on the RHS of Eq. (5.45), however, cannot be considered as the Fréchet derivative of $\Delta(v)$ since the density function $\rho(y)$ is evaluated at $\xi(x)$ and not $v(x)$. As such, we rewrite Eq. (5.45) as follows,

$$\begin{aligned} \Delta(v+h) - \Delta(v) &= -2\int_X [P(u(x)) - P(v(x))][\rho(v(x)) + \rho(\xi(x)) - \rho(v(x))]h(x)dx \\ &\quad + \int_X \rho(\xi(x))^2 h(x)^2 dx \\ &= -2\int_X [P(u(x)) - P(v(x))]\rho(v(x))h(x)dx \\ &\quad - 2\int_X [P(u(x)) - P(v(x))][\rho(\xi(x)) - \rho(v(x))]h(x)dx + \int_X \rho(\xi(x))^2 h(x)^2 dx. \end{aligned} \tag{5.46}$$

It now appears that the first term on the RHS of Eq. (5.46) could be considered as the Fréchet derivative of $\Delta(v)$, i.e.,

$$\Delta'(v)h = -2\int_X [P(u(x)) - P(v(x))]\rho(v(x))h(x)dx, \tag{5.47}$$

At first glance, the second term on the RHS of Eq. (5.46) seems linear in h; However, we shall show that it goes to zero faster than the linear form of h when h tends to zero, because $\rho(y)$ is continuous and $\xi \to v$ as $h \to 0$.

Let us first rewrite Eq. (5.46) as follows, based on the assumption in Eq. (5.47),

$$\begin{aligned}\Delta(v+h) - \Delta(v) - \Delta'(v)h = -2\int_X [P(u(x)) - P(v(x))][\rho(\xi(x)) - \rho(v(x))]h(x)dx \\ + \int_X \rho(\xi(x))^2 h(x)^2 dx.\end{aligned} \tag{5.48}$$

Since Δ is a functional, i.e., $\Delta : \mathcal{F}(X) \to \mathbb{R}$, we may take absolute values of Eq. (5.48) and obtain

$$\begin{aligned}|\Delta(v+h) - \Delta(v) - \Delta'(v)h| \le 2\int_X |[P(u(x)) - P(v(x))][\rho(\xi(x)) - \rho(v(x))]h(x)|dx \\ + \int_X \rho(\xi(x))^2 h(x)^2 dx.\end{aligned} \tag{5.49}$$

In fact, the second integral on the RHS of the above equation can be bounded in a number of ways. One is to use the facts that $\rho(y)$ is continuous and $\mathbb{R}_g = [A, B]$ is compact, which imply that there exists an $M > 0$ such that

$$|\rho(y)| \le M \quad \forall y \in \mathbb{R}_g, \tag{5.50}$$

which, in turn, implies that

$$|\rho(\xi(x))| \le M \quad \text{a.e. } x \in X. \tag{5.51}$$

Then, we have

$$\int_X \rho(\xi(x))^2 h(x)^2 dx \le M^2 \|h\|_2^2. \tag{5.52}$$

With regard to the first integral on the RHS of Eq. (5.49), we first examine the term, i.e., $\rho(\xi(x)) - \rho(v(x))$, in the integrand. Recall that for a.e. $x \in X$, $\xi(x)$ lies between $v(x)$ and $v(x) + h(x)$, which indicates that

$$|\xi(x) - v(x)| \le |h(x)| \quad \text{a.e. } x \in X. \tag{5.53}$$

Note that $\rho(y)$ is uniformly continuous on $\mathbb{R}_g$, which implies that for any $\epsilon > 0$, there exists a $\delta > 0$ so that

$$|\rho(y_1) - \rho(y_1)| < \epsilon \text{ for all } y_1, y_2 \in \mathbb{R}_g \text{ such that } |y_1 - y_2| < \delta. \tag{5.54}$$

Now, from Eqs. (5.53) and (5.54), if $h \in H_n$ satisfies the condition $|h(x)| < \delta$ for a.e. $x \in X$, then

$$|\rho(\xi(x)) - \rho(v(x))| < \epsilon \quad \text{a.e.} \quad x \in X. \tag{5.55}$$

The first integral on the RHS of Eq. (5.49) can then be bounded as follows,

$$\int_X |[P(u(x)) - P(v(x))][\rho(\xi(x)) - \rho(v(x))]h(x)|dx \leq \epsilon \int_X |P(u(x)) - P(v(x))||h(x)|dx, \tag{5.56}$$

which implies that

$$\int_X |[P(u(x)) - P(v(x))][\rho(\xi(x)) - \rho(v(x))]h(x)|dx \leq \epsilon D_2(u, v; \nu)\|h\|_2, \tag{5.57}$$

by imposing Cauchy-Schwartz inequality on the RHS of Eq. (5.56), where $D_2(u, v; \nu)$ has been shown in Eq. (5.2). Employing the results of Eqs. (5.52) and (5.41) on Eq. (5.49), we obtain the following result,

$$\frac{|\Delta(v + h) - \Delta(v) - D\Delta(v)h|}{\|h\|_2} \leq M^2\|h\|_2 + \epsilon D_2(u, v; \nu). \tag{5.58}$$

Since $|h(x)| < \delta$ for a.e. $x \in X$, $\|h\|_2 < \delta\sqrt{L}$, where L represents the "length" of X. By letting $\epsilon \to 0^+$ and considering $\|h\|_2 < \min\{\delta\sqrt{L}, \epsilon\}$, the RHS can be made arbitrarily small, which proves that the Fréchet derivative of $\Delta(v)$ is given in Eq. (5.47).

Technically speaking, the denominator of the LHS in Eq. (5.58) should employ the true distance function D_2 between the functions $v + h$ and v, i.e.,

$$D_2(v + h, v; \nu) = \left[\int_X [P(v(x) + h(x)) - P(v(x))]^2 \, dx\right]^{1/2}, \tag{5.59}$$

rather than $\|h\|_2$. However, in the case that the density function $\rho(y)$ is continuous on $\mathbb{R}_g$, we shall show the D_2 metric is equivalent to the usual L^2 metric. As such, Eq. (5.59) could be rewritten in terms of the usual L^2 metric, establishing the existence of the Fréchet derivative of $D_2(u, v; \nu)$. Hence, $D_2(u, v; \nu)$ is Fréchet differentiable on the function space S_n (i.e., Eq. (5.5)).

In fact, the proof of equivalence of D_2 and L^2 metrics can be considered as a generalization of the one given in Section 3.2.3 for the particular case of Weber-based density functions having the form $\rho_a(y) = 1/y^a$ for $0 \leq a \leq 1$ defined over the positive greyscale range space $\mathbb{R}_g = [A, B]$ where $0 < A, B < \infty$. In Section 3.2.3, we proved the equivalence

of L^2 metric and Weber-based metrics constructed by the nonuniform measures defined by the density functions of the form $\rho_a(y) = 1/y^a$ for $0 \leq a \leq 1$ over the positive greyscale range space $\mathbb{R}_g = [A, B]$ where $0 < A, B < \infty$.

Here, we may consider the more general case in which $\mathbb{R}_g = [A, B] \subset \mathbb{R}$. Once again, we assume that the measure ν supported on $\mathbb{R}_g$ is defined by a density function $\rho(y) > 0$ which is continuous on $\mathbb{R}_g$. The metric $D_2(u, v; \nu)$ on $\mathcal{F}([a, b])$ constructed by the intensity-based measure ν is defined in Eq. (5.2), where P is the antiderivative of ρ, i.e., $P'(y) = \rho(y)$. For a.e. $x \in [a, b]$, applying the Mean Value Theorem to the function $P(y)$ on $[A, B]$ yields,

$$\begin{aligned} P(u(x)) - P(v(x)) &= P'(c)(u(x) - v(x)) \\ &= \rho(c)(u(x) - v(x)), \end{aligned} \tag{5.60}$$

where c lies between $u(x)$ and $v(x)$. Taking absolute values, and noting that $A < c < B$, we obtain the following set of inequalities for a.e., $x \in [a, b]$,

$$\rho_{\min}|u(x) - v(x)| \leq |P(u(x)) - P(v(x))| \leq \rho_{\max}|u(x) - v(x)|, \tag{5.61}$$

where

$$\rho_{\min} = \min_{A \leq y \leq B} \rho(y) > 0, \quad \rho_{\max} = \max_{A \leq y \leq B} \rho(y). \tag{5.62}$$

By squaring all terms of Eq. (5.61), integrating over $[a, b]$ and taking square roots, we obtain the final result as follows,

$$\rho_{\min} d_2(u, v) \leq D_2(u, v; \nu) \leq \rho_{\max} d_2(u, v), \tag{5.63}$$

where

$$d_2(u, v) = \left[\int_a^b [u(x) - v(x)]^2 dx \right]^{1/2}. \tag{5.64}$$

As a result, the equivalence of the usual L^2 and D_2 metrics has been established. Hence, the effectiveness of the Fréchet derivative of $\Delta(v)$ in Eq. (5.47) has been validated. Finally, let us recall the meaning of the Fréchet derivative of $\Delta(v)$ as given in Eq. (5.47): $\Delta'(v)h$ is the instantaneous rate of change of the functional $\Delta(v)$ at the point $v \in \mathcal{F}([a, b])$ in the direction h.

The Minimization of Generalized Weber-based Metrics using the Stationarity Conditions

Motivated by the results of Fréchet derivative of generalized Weber-based metrics, one may wonder whether the stationary points $v_n \in S_n$ are the minimizers of the distance functional

$\Delta(v)$ with reference to Eq. (5.30). In other words, if

$$\Delta'(v_n)h = 0 \quad \forall h \in H_n \quad \text{such that} \quad v_n + h \in S_n, \tag{5.65}$$

does it follow that

$$v_n = \arg \min_{v \in S_n} \Delta(v), \tag{5.66}$$

or

$$\Delta(v_n + h) \geq \Delta(v_n)? \tag{5.67}$$

If we recall the case that the density function $\rho(y) = 1$ implying the Lebesgue measure, the answer is "Yes," cf. Eq. (5.39). From Eq. (5.48), if $v_n \in S_n$ such that $\Delta'(v_n) = 0$, then for any $h \in H_n$ such that $v_n + h \in S_n$,

$$\begin{aligned} \Delta(v_n + h) - \Delta(v_n) = -2 \int_X [P(u(x)) - P(v_n(x))][\rho(\xi(x)) - \rho(v_n(x))]h(x)dx \\ + \int_X \rho(\xi(x))^2 h(x)^2 dx. \end{aligned} \tag{5.68}$$

where, for a.e. $x \in X$, $\xi(x)$ lies between $v_n(x)$ and $v_n(x)+h(x)$. If we assume that $\|h\|_2 \neq 0$, the second integral on the RHS of the above equation is certainly positive. Moreover, it follows from $v_n \in S_n$ is a stationary point of $\Delta(v)$ that the above equation becomes

$$\begin{aligned} \Delta(v_n + h) - \Delta(v_n) = -2 \int_X [P(u(x)) - P(v_n(x))]\rho(\xi(x))h(x)dx \\ + \int_X \rho(\xi(x))^2 h(x)^2 dx, \end{aligned} \tag{5.69}$$

which was actually derived earlier as Eq. (5.45).

From Eq. (5.69) it follows that if $\|h\|_2$ is small, then ξ is close to v, which would imply that first integral on the RHS of the above equation is close to zero in value. There would then be the chance that the net value of the RHS of Eq. (5.69) would be positive. Indeed, the second integral in Eq. (5.69) can be bounded below as follows,

$$\int_X \rho(\xi(x))^2 h(x)^2 dx \geq \rho_{\min}^2 ||h||_2^2, \tag{5.70}$$

where $|| \cdot ||_2$ denotes the usual L^2 norm. We now try to obtain an upper bound to the first integral, which will then give the most negative possible contribution to the RHS of Eq. (5.68). One such bound can be found if we assume that $\rho(y)$ is not only uniform

continuous but uniformly Lipschitz on $\mathbb{R}_g$ – which is not an unreasonable assumption. In this case there exists a $K \geq 0$ such that

$$|\rho(\xi(x)) - \rho(v(x))| \leq K|\xi(x) - v(x)| \leq K|h(x)| \quad \text{a.e.} \quad x \in X. \tag{5.71}$$

For the case of signal functions approximation, let us assume that $X = [a, b]$. It then follows from the above equation that the first integral on the RHS of Eq. (5.68) can then be bounded as follows,

$$\begin{aligned}\int_X |[P(u(x)) - P(v_n(x))][\rho(\xi(x)) - \rho(v_n(x))]h(x)|dx &\leq K \int_X |P(u(x)) - P(v_n(x))||h(x)|^2 dx \\ &\leq K(b-a)\nu(\mathbb{R}_g)||h||_2^2.\end{aligned} \tag{5.72}$$

Note that the ν-measure of an interval $[y_1, y_2] \subseteq \mathbb{R}_g$ is given by

$$\nu([y_1, y_2]) = \int_{y_1}^{y_2} \rho(y)dy = P(y_2) - P(y_1) \leq \nu(\mathbb{R}_g), \tag{5.73}$$

where $\nu(\mathbb{R}_g)$ is the whole measure of the greyscale range space $\mathbb{R}_g$. Consequently, it follows from $u(x), v(x) \in \mathbb{R}_g$ for a.e. $x \in [a, b]$ that the second line of Eq. (5.72) can be obtained. If we combine the two bounds showed in Eqs. (5.70) and (5.72), we have, from Eq. (5.69), that

$$D(v_n + h) - D(v_n) \geq [\rho_{\min}^2 - 2K(b-a)\nu(\mathbb{R}_g)]||h||_2^2. \tag{5.74}$$

The RHS is guaranteed to be positive if

$$\rho_{\min}^2 > 2K(b-a)\nu(\mathbb{R}_g). \tag{5.75}$$

Note that we cannot ensure that this inequality is satisfied by simply increasing $\rho_{\min}$, i.e., by multiplying $\rho(y)$ by a constant $C > 0$. By doing this, we also multiply $\nu(\mathbb{R}_g)$, as well as the Lipschitz constant K, by C. One thing that can be done immediately is to let $a = 0$ and $b = 1$ so that $b - a = 1$, i.e., a rescaling of the x-coordinate. Now, the inequality becomes

$$\rho_{\min}^2 > 2K\nu(\mathbb{R}_g). \tag{5.76}$$

Note that in the case of Lebesgue measure, i.e., $\rho(y) = C$, the Lipschitz constant $K = 0$, so the above inequality is satisfied. In this case, we know that v_n is a minimizer. From the asymptotic results determined in Chapter 4, we also have that

- For $0 < a < 1$:

$$D_a(u, v) = \int_X [u(x)^{-a+1} - (v(x))^{-a+1}]^2 dx. \tag{5.77}$$

- For $a = 1$:

$$D_1(u, v) = \int_X [\log u(x) - \log(v(x))]^2 dx. \tag{5.78}$$

Now, let us consider the following cases where the density functions, i.e., $\rho(y)$, are positive-valued and decreasing. In fact, these cases correspond to Weber-based intensity measures on the bounded greyscale intensity range, i.e., $\mathbb{R}_g = [A, B]$, so that

$$\int_A^B \rho(y)dy = \nu(\mathbb{R}_g). \tag{5.79}$$

Case 1: Squared Generalized Weber-based L^2 Metrics

For the case that $0 < a < 1$, $\rho(y) = 1/y^a$ on $[A, B]$, and the Lipschitz constant $K = a/A^{a+1}$. Note that these generalized Weber-based L^2 metrics are introduced in Eq. (5.77). From Eq. (5.79), we obtain that

$$\nu(\mathbb{R}_g) = \int_A^B \frac{1}{y^a} dy = \frac{1}{1-a}\left[B^{-a+1} - A^{-a+1}\right]. \tag{5.80}$$

Plugging all the facts including $\rho_{\min} = 1/B^a$, $K = a/A^{a+1}$ and Eq. (5.80) into Eq. (5.76) yields

$$\frac{1}{B^{2a}} > \frac{2a}{(1-a)A^{a+1}}\left[B^{-a+1} - A^{-a+1}\right]. \tag{5.81}$$

After some manipulation, we have

$$\frac{2a}{(1-a)}\frac{B^{a+1}}{A^{a+1}} - \frac{B^{2a}}{A^{2a}} - 1 < 0. \tag{5.82}$$

If we let $r = B/A$, the above equation becomes

$$\frac{2a}{(1-a)} r^{a+1} - r^{2a} - 1 < 0. \tag{5.83}$$

If we consider the LHS of Eq. (5.83) as a new function,

$$R(r) = \frac{2a}{(1-a)} r^{a+1} - r^{2a} - 1, \tag{5.84}$$

it is clear that

$$\begin{aligned} R'(r) &= \frac{2a(a+1)}{(1-a)} r^a - 2ar^{2a-1} \\ &= 2ar^a \left[\frac{1+a}{1-a} - r^{a-1} \right] \\ &= 2ar^a \left[1 + \frac{2a}{1-a} - r^{a-1} \right]. \end{aligned} \tag{5.85}$$

Since $r > 1$ and $0 < a < 1$, $r^{a-1} < 1$, which implies that $R'(r) > 0$ for all $r > 1$. As such, we obtain that

$$R(r) > R(1) = \frac{2a}{(1-a)} - 2 = \frac{4a-2}{1-a}. \tag{5.86}$$

- If $\frac{1}{2} \leq a < 1$, then $R(r) > R(1) \geq 0$, which suggests that there is no solution satisfying Eq. (5.83). Therefore, for any $\frac{1}{2} \leq a < 1$, the stationary points of the distance functions $D_a(u, v)$, i.e., Eq. (5.77), cannot be guaranteed to be the local minimizers of $D_a(u, v)$.

- If $0 < a < \frac{1}{2}$, we have $R(1) < 0$. The solution of Eq. (5.83) is as follows,

$$1 < r < \bar{r}, \tag{5.87}$$

  where $\bar{r} > 1$ is the zero point (if exist) of $R(r)$ in Eq. (5.84), i.e. $R(\bar{r}) = 0$. Notice that if there exists such a point $\bar{r}$, i.e., there exists a point $r_1 > 1$ such that $R(r_1) > 0$, $\bar{r}$ is the only zero point of $R(r)$ for $r \in (1, \infty)$ since $R'(r) > 0$ for all $r > 1$. On the other hand, if $R(r) < 0$ for all $r \in (1, \infty)$,then there is no zero point with respect to function $R(r)$, so $\bar{r}$ becomes ∞. In practice, the existence and location of $\bar{r}$ can be justified and "detected" by some numerical techniques such as Newton's method, Secant method and Brent's method, we then have a range of the ratio between A and B, i.e., Eq. (5.87), such that Eq. (5.76) is satisfied.

In what follows, let us test a simple example with $a = 1/4$. It follows from Eqs. (5.84) and (5.85) that

$$R(r) = \frac{2}{3} r^{\frac{5}{4}} - r^{\frac{1}{2}} - 1, \tag{5.88}$$

and

$$R'(r) = \frac{5}{6} r^{\frac{1}{4}} - \frac{1}{2} r^{-\frac{1}{2}}. \tag{5.89}$$

We aim to implement Newton's method to approximate the root of Eq. (5.88). With reference to Eqs. (5.88) and (5.89), the iteration function of Newton's method is given by

$$\begin{aligned} r_{n+1} &= r_n - \frac{R(r_n)}{R'(r_n)} \\ &= r_n - \frac{4r_n^{\frac{5}{4}} - 6r_n^{\frac{1}{2}} - 6}{5r_n^{\frac{1}{4}} - 3r_n^{-\frac{1}{2}}}, \end{aligned} \tag{5.90}$$

where $n = 0, 1, \cdots, N$, and N denotes total number of iterations. Using MATLAB with our initial guess of $r_0 = 5$, we can produce the following values:

$$\begin{aligned} r_1 &= 3.2900747461, \\ r_2 &= 3.1245081676, \\ r_3 &= 3.1223592408, \\ r_4 &= 3.1223588658, \\ r_5 &= 3.1223588658. \end{aligned} \tag{5.91}$$

Accordingly, we also have

$$\begin{aligned} R(r_1) &= 1.7484279599, \\ R(r_2) &= 0.1401807905, \\ R(r_3) &= 0.0017730144, \\ R(r_4) &= 3.0927155718 \times 10^{-7}, \\ R(r_5) &= 9.4210307000 \times 10^{-15}. \end{aligned} \tag{5.92}$$

Additionally, $R(4) \approx 0.77123 > 0$, which automatically shows there exists a unique root of $R(r)$ in Eq. (5.88). As a result, $\bar{r} \approx 3.1223588658$, which is the upper bound of r, i.e., B/A, so that Eq. (5.88) is negative for $r \in (1, \bar{r})$. This indeed implies that for $a = 1/4$ and $1 < B/A < 3.1223588658$, the stationary points of the distance function $D_{\frac{1}{4}}(u, v)$, referring to Eq. (5.77), are minimizers of $D_{\frac{1}{4}}(u, v)$. Therefore, for any given target function $u \in \mathcal{F}(X)$ (i.e., Eq. (5.4)), if the lower and upper bounds (i.e., A and B) of greyscale range $\mathbb{R}_g$ satisfies the following relation,

$$1 < \frac{B}{A} < 3.1223588658, \tag{5.93}$$

and there exists a v_n such that

$$D'_{\frac{1}{4}}(u, v_n)h = 0 \quad \forall h \in H_n \quad \text{such that} \quad v_n + h \in S_n, \tag{5.94}$$

then v_n is consequently considered as a minimizer of the distance function $D_{\frac{1}{4}}(u, v)$.

Case 2: Squared Logarithmic L^2 Metric

As a special case of $a = 1$ corresponding to standard Weberized metric (i.e., Eq. (5.78)), $\rho(y) = 1/y$ on $[A, B]$, and the Lipschitz constant $K = 1/A^2$. It then follows from Eq. (5.79) that

$$\nu(\mathbb{R}_g) = \int_A^B \frac{1}{y} dy = \ln\left(\frac{B}{A}\right). \tag{5.95}$$

Consequently, Eq. (5.78) gives to

$$\frac{1}{B^2} > \frac{2}{A^2} \ln\left(\frac{B}{A}\right), \tag{5.96}$$

since $\rho_{\min} = 1/B$. After a little manipulation, we have

$$\ln r - \frac{1}{2r^2} < 0, \tag{5.97}$$

by the change of variable, i.e., $r = B/A$. Note that, in general, $B > A$, which means that $r \in (1, \infty)$. Let us consider a new function as follows,

$$R(r) = \ln r - \frac{1}{2r^2}. \tag{5.98}$$

It is straightforward to see that $R'(r) > 0$ for all $r \in (1, \infty)$, which implies that

$$R(r) > R(1) = -\frac{1}{2}. \tag{5.99}$$

In addition, $R(e) = 1 - 1/2e^2 > 0$. Hence, there exists a unique $\bar{r} \in (1, \infty)$ such that $R(\bar{r}) = 0$. To perform the iteration of Newton's method, we need to know the derivative $R'(r) = 1/r + 1/r^3$ so that

$$r_{n+1} = r_n - \frac{2r_n^3 \ln r_n - r_n}{2r_n^2 + 2}. \tag{5.100}$$

Starting at our initial guess of $r_0 = 3$, we obtain the following numerical results,

$$\begin{aligned} r_1 &= 0.1837468205, \\ r_2 &= 0.2827868501, \\ r_3 &= 0.4401584188, \\ r_4 &= 0.6831414991, \\ r_5 &= 0.9988570150, \\ r_6 &= 1.2494273641, \\ r_7 &= 1.3237630678, \\ r_8 &= 1.3278530806, \\ r_9 &= 1.3278640119, \\ r_{10} &= 1.3278640119, \end{aligned} \tag{5.101}$$

Note that $R(r_{10}) = 6.1062266354 \times 10^{-6}$, which is very close to zero. As such, $\bar{r} \approx r_{15} = 1.32786401199$, which is the upper bound of r, i.e., B/A, so that Eq. (5.98) is negative for $r \in (1, \bar{r})$. Therefore, for squared standard Weberized metric, i.e., Eq. (5.78), if

$$1 < \frac{B}{A} < 1.3278640119, \tag{5.102}$$

the stationary points of $D_1(u, v)$ are also the minimizers of $D_1(u, v)$. Note that even if A and B do not satisfy the above relation, a stationary point could possibly be an optimal solution that minimizes standard Weberized metric $D_1(u, v)$.

Admittedly, the results in Eqs. (5.93) and (5.102) are not very encouraging due to the fact that A/B is bounded in a very small "region", in particular, as shown in Eq. (5.102). Nevertheless, Eqs. (5.93) and (5.102) may provide some help with best approximation problems in minimizing the generalized Weber-based metrics ($0 < a < 1/2$) and standard Weberized metric ($a = 1$) from computational perspective.

5.3 The Fréchet Derivative of SSIM between Two Image Functions

5.3.1 Introduction

In [17], the authors examined the best approximation problem using structural similarity (SSIM): Given a reference function u, find the approximation $v \in \mathcal{F}(X)$, where $\mathcal{F}(X)$

(defined in Eq. (5.4)) denotes a suitable class of positive-valued functions on X, such that the SSIM, which measures the perceptual difference, between u and v is maximized. Without loss of generality, we assume that u and v are integrable functions on the interval $[a, b]$, i.e., $X = [a, b]$. Ignoring the numerical stability constants, the SSIM function between u and v is defined as follows,

$$S(u, v) = \frac{4\bar{u}\bar{v}\sigma_{uv}}{(\bar{u}^2 + \bar{v}^2)(\sigma_u^2 + \sigma_v^2)}, \tag{5.103}$$

where

$$\bar{v} = \frac{1}{b-a}\int_a^b v(x)dx = \langle v \rangle, \tag{5.104}$$

$$\sigma_v^2 = \frac{1}{b-a}\int_a^b (v - \bar{v})^2 dx, \tag{5.105}$$

representing the mean and variance of the function v over the interval $[a, b]$, respectively. Similar expressions hold for u. Moreover,

$$\sigma_{uv} = \frac{1}{b-a}\int_a^b (u - \bar{u})(v - \bar{v})dx, \tag{5.106}$$

which denotes the covariance between functions u and v. In this section, let us consider the approximations afforded by a family of complete orthonormal functions $\{\phi_k\}_{k=1}^{\infty}$ on $[a, b]$, i.e.,

$$v_n(x) = \sum_{k=1}^{n} c_k \phi_k(x), \tag{5.107}$$

where

$$\langle \phi_k, \phi_l \rangle = \int_a^b \phi_k(x)\phi_l(x)dx = \delta_{kl}. \tag{5.108}$$

In this case, $\mathcal{F}(X)$ becomes $\text{span}\{\phi_k\}_{k=1}^{n}$ on $[a, b]$. To find the best approximation of the reference function u, we shall work out a set of optimal coefficients $\mathbf{c} = (c_1, \cdots, c_n)$ that maximizes $S(u, v_n)$. As such, we rewrite Eq. (5.103) in terms of the parameters $\mathbf{c}$,

$$S(\mathbf{c}) = \frac{4\bar{u}\bar{v}_n(\mathbf{c})\sigma_{uv_n}(\mathbf{c})}{(\bar{u}^2 + \bar{v}_n(\mathbf{c})^2)(\sigma_u^2 + \sigma_{v_n}(\mathbf{c})^2)}. \tag{5.109}$$

Note that, as the "input" information, the terms $\bar{u}$ and σ_u^2 are specific constants which are not dependent upon $\mathbf{c}$. Then, we employ logarithmic differentiation to simplify the calculations on computing the gradient $\nabla_{\mathbf{c}} S(u, v_n(\mathbf{c}))$, i.e.,

$$\log S = \log(4\bar{u}) + \log(\bar{v}_n(\mathbf{c})) + \log(\sigma_{uv_n}(\mathbf{c})) - \log(\bar{u}^2 + \bar{v}_n(\mathbf{c})^2) - \log(\sigma_u^2 + \sigma_{v_n}^2(\mathbf{c})), \tag{5.110}$$

and

$$\frac{1}{S}\frac{\partial S}{\partial c_p} = \frac{1}{\bar{v}_n}\frac{\partial \bar{v}_n}{\partial c_p} + \frac{1}{\sigma_{uv_n}}\frac{\partial \sigma_{uv_n}}{\partial c_p} - \frac{1}{\bar{u}^2+\bar{v}_n^2}\frac{\partial \bar{v}_n^2}{\partial c_p} - \frac{1}{\sigma_u^2+\sigma_{v_n}^2}\frac{\partial \sigma_{v_n}^2}{\partial c_p}, \tag{5.111}$$

where $p \in \{1, \cdots, n\}$. After a little manipulation on Eq. (5.96), we obtain,

$$\frac{\partial S}{\partial c_p} = S\left[\frac{1}{\bar{v}_n}\frac{\partial \bar{v}_n}{\partial c_p} + \frac{1}{\sigma_{uv_n}}\frac{\partial \sigma_{uv_n}}{\partial c_p} - \frac{1}{\bar{u}^2+\bar{v}_n^2}\frac{\partial \bar{v}_n^2}{\partial c_p} - \frac{1}{\sigma_u^2+\sigma_{v_n}^2}\frac{\partial \sigma_{v_n}^2}{\partial c_p}\right]. \tag{5.112}$$

Now, it remains to compute the partial derivatives on the RHS of Eq. (5.112). Indeed, from Eqs. (5.104), (5.105), (5.106) and (5.107), we can obtain that

$$\frac{\partial \bar{v}_n}{\partial c_p} = \frac{\partial}{\partial c_p}\left[\sum_{k=1}^{n} c_k\langle\phi_k(x)\rangle\right] = \langle\phi_p(x)\rangle, \tag{5.113}$$

and

$$\begin{aligned}
\frac{\partial \sigma_{uv_n}}{\partial c_p} &= \frac{\partial}{\partial c_p}\left[\frac{1}{b-a}\int_a^b (u-\bar{u})\left[\sum_{k=1}^{n} c_k\phi_k(x) - \sum_{k=1}^{n} c_k\langle\phi_k(x)\rangle\right]dx\right] \\
&= \frac{1}{b-a}\int_a^b (u-\bar{u})(\phi_p(x) - \langle\phi_p(x)\rangle)dx \\
&= \frac{1}{b-a}[a_p - \langle u\rangle\langle\phi_p(x)\rangle - \langle u\rangle\langle\phi_p(x)\rangle + \langle u\rangle\langle\phi_p(x)\rangle(b-a)] \\
&= \frac{1}{b-a}[a_p + (b-a-2)\langle u\rangle\langle\phi_p(x)\rangle],
\end{aligned} \tag{5.114}$$

where $a_p = \langle u, \phi_p(x)\rangle$, namely, "Fourier coefficients", derived from best usual L^2 approximation using the set of orthonormal functions $\{\phi_k\}_{k=1}^{\infty}$ on $[a, b]$. Moreover, since

$$\begin{aligned}
\sigma_{v_n}^2 &= \frac{1}{b-a}\int_a^b (v_n - \bar{v}_n)(v_n - \bar{v}_n)dx \\
&= \frac{1}{b-a}\int_a^b (v_n^2 - 2v_n\bar{v}_n + \bar{v}_n^2)dx \\
&= \frac{1}{b-a}[\langle v_n^2\rangle + (b-a-2)\bar{v}_n^2],
\end{aligned} \tag{5.115}$$

it follows from Eqs. (5.113) and (5.115) that

$$\begin{aligned}
\frac{\partial \sigma_{v_n}^2}{\partial c_p} &= \frac{1}{b-a}[2\langle v_n, \phi_p(x)\rangle + 2(b-a-2)\bar{v}_n\langle\phi_p(x)\rangle] \\
&= \frac{1}{b-a}\left[2\sum_{k=1}^{n} c_k\langle\phi_k(x), \phi_p(x)\rangle + 2(b-a-2)\bar{v}_n\langle\phi_p(x)\rangle\right].
\end{aligned} \tag{5.116}$$

From Eq. (5.108), the above equation becomes,

$$\frac{\partial \sigma_{v_n}^2}{\partial c_p} = \frac{1}{b-a}[2c_p + 2(b-a-2)\bar{v}_n\langle\phi_p(x)\rangle]. \tag{5.117}$$

Finally, imposing the Chain Rule and employing Eq. (5.113) yields,

$$\frac{\partial \bar{v}_n^2}{\partial c_p} = 2\bar{v}_n\langle\phi_p(x)\rangle. \tag{5.118}$$

Hence, by substituting these results into Eq. (5.112), we obtain

$$\frac{\partial S}{\partial c_p} = S\left[\frac{\langle\phi_p\rangle}{\bar{v}_n} + \frac{a_p + (b-a-2)\langle u\rangle\langle\phi_p\rangle}{\sigma_{uv_n}(b-a)} - \frac{2\bar{v}_n\langle\phi_p\rangle}{\bar{u}^2+\bar{v}_n^2} - \frac{2c_p + 2(b-a-2)\bar{v}_n\langle\phi_p\rangle}{(\sigma_u^2+\sigma_{v_n}^2)(b-a)}\right], \tag{5.119}$$

where $p = 1, \cdots, n$. In general, the RHS of Eq. (5.119) is a complicated nonlinear system of equations in the unknown coefficients c_p and a solution of the equations $\partial S/\partial c_k = 0$ is intractable. One may propose to employ the Gauss-Newton strategy, an efficient gradient-based optimization algorithm [42], to locate the maximum points of $S(u, v(\mathbf{c}))$. In many applications of image processing, given the oscillatory nature of the basis $\{\phi_k\}_{k=1}^n$, many of the mean values $\langle\phi_k\rangle$ for a given orthonormal set may be zero. In this case, the above expression simplifies greatly:

$$\frac{\partial S}{\partial c_p} = S\left[\frac{a_p}{\sigma_{uv_n}(b-a)} - \frac{2c_p}{(\sigma_u^2+\sigma_{v_n}^2)(b-a)}\right]. \tag{5.120}$$

That being said, the vanishing of the partial derivative, i.e., $\partial S/\partial c_p$, implies that

$$a_p(\sigma_u^2 + \sigma_{v_n}^2) = 2\sigma_{uv_n}c_p, \tag{5.121}$$

which is a quadratic form in terms of the unknown coefficients c_p for $p = 1, \cdots, n$. If we look at the above equation more carefully, we shall realize that the stationarity equation (i.e., $\partial S/\partial c_p = 0$) is not depend upon the constants a and b, which is not surprising in a sense of the best approximation theory, because, intuitively, the optimal solution v_n satisfying Eq. (5.121) is supposed to be restricted by the "size" of the space of functions $\mathcal{F}(X)$ rather than the length of the interval $X = [a, b]$ (i.e., $b - a$).

5.3.2 Consequences of Stationarity and the Fréchet Derivative of SSIM

Previously, we investigated the stationary point of SSIM, which involves the partial derivatives with respect to the set of unknown parameters $\mathbf{c} = (c_1, \cdots, c_n)$. As indicated before,

$v_n \in \mathcal{F}([a,b]) \subseteq L^2[a,b]$ is the approximation to u, where a, b are non-negative constants, and a set of functions ϕ_k for $k \in \{1, \cdots, n\}$ form an orthonormal basis on $[a,b]$. Recall that we considered the function v_n described by Eq. (5.107) (i.e., $v_n \in S_n$) as an approximation to a given function $u \in \mathcal{F}([a,b])$. Here we consider the best approximation of u in terms of SSIM, i.e., the function v_n which maximizes the SSIM between u and v_n, denoted as $S(u, v_n)$ in Eq. (5.103). From Eq. (5.119), the following stationarity conditions must be satisfied,

$$\left[\frac{\langle\phi_p\rangle}{\bar{v}_n} + \frac{a_p + (b-a-2)\langle u\rangle\langle\phi_p\rangle}{\sigma_{uv_n}(b-a)} - \frac{2\bar{v}_n\langle\phi_p\rangle}{\bar{u}^2+\bar{v}_n^2} - \frac{2c_p + 2(b-a-2)\bar{v}_n\langle\phi_p\rangle}{(\sigma_u^2+\sigma_{v_n}^2)(b-a)}\right] = 0, \quad (5.122)$$

for $p = 1, \cdots, n$, since we have assumed $S(u, v_n) \neq 0$. Note that, as before, a_p denotes the "Fourier coefficients".

Motivated by the previous work on intensity-based measures, i.e., Section 5.1.2, we shall now perform a series of operations on Eq. (5.106) to further analyze mathematical properties of SSIM between u and v_n. First of all, consider any function $h_n \in \mathcal{F}([a,b])$, defined as follows, be

$$h_n(x) = \sum_{k=1}^{n} d_k\phi_k(x). \quad (5.123)$$

For each $p \in \{1, \cdots, n\}$, multiply both sides of Eq.(5.106) by the coefficient d_p,

$$\left[\frac{\langle\phi_p d_p\rangle}{\bar{v}_n} + \frac{a_p d_p + (b-a-2)\langle u\rangle\langle\phi_p d_p\rangle}{\sigma_{uv_n}(b-a)} - \frac{2\bar{v}_n\langle\phi_p d_p\rangle}{\bar{u}^2+\bar{v}_n^2} - \frac{2c_p d_p + 2(b-a-2)\bar{v}_n\langle\phi_p d_p\rangle}{(\sigma_u^2+\sigma_{v_n}^2)(b-a)}\right] = 0. \quad (5.124)$$

Now sum over all $p \in \{1, \cdots, n\}$, we have,

$$\left[\frac{\langle h_n\rangle}{\bar{v}_n} + \frac{\sum_{p=1}^n a_p d_p + (b-a-2)\langle u\rangle\langle h_n\rangle}{\sigma_{uv_n}(b-a)} - \frac{2\bar{v}_n\langle h_n\rangle}{\bar{u}^2+\bar{v}_n^2} - \frac{2\sum_{p=1}^n c_p d_p + 2(b-a-2)\bar{v}_n\langle h_n\rangle}{(\sigma_u^2+\sigma_{v_n}^2)(b-a)}\right] = 0, \quad (5.125)$$

which may be rewritten as follows,

$$\frac{\langle h_n\rangle}{\bar{v}_n} + \frac{[\langle u, h_n\rangle + (b-a-2)\bar{u}\bar{h}_n]}{\sigma_{uv_n}(b-a)} - \frac{2\bar{v}_n\bar{h}_n}{\bar{u}^2+\bar{v}_n^2} - \frac{[2\langle v_n, h_n\rangle + 2(b-a-2)\bar{v}_n\bar{h}_n]}{(\sigma_u^2+\sigma_{v_n}^2)(b-a)} = 0. \quad (5.126)$$

Note that this result is valid for all functions $h_n \in \mathcal{F}([a,b])$. Moreover, the LHS of the above equation has the form of a linear operator acting on h_n, since the mean operators (e.g., $\langle u, h_n\rangle$) are linear in their arguments. This once again suggests that the LHS is a Fréchet derivative. That being said, if the structural similarity between the fixed target

function u and the variable approximation v_n is now considered to be a functional in terms of v_n, i.e.,

$$S(v) = \frac{4\bar{u}\bar{v}\sigma_{uv}}{(\bar{u}^2+\bar{v}^2)(\sigma_u^2+\sigma_v^2)}. \tag{5.127}$$

then we conjecture that the Fréchet derivative of the functional $S(v_n)$ is given by

$$S'(v_n)h = S(v_n)\left[\frac{\langle h_n\rangle}{\bar{v}_n} + \frac{\langle u, h_n\rangle + (b-a-2)\bar{u}\bar{h}_n}{\sigma_{uv_n}(b-a)} - \frac{2\bar{v}_n\bar{h}_n}{\bar{u}^2+\bar{v}_n^2} - \frac{2\langle v_n, h_n\rangle + 2(b-a-2)\bar{v}_n\bar{h}_n}{(\sigma_u^2+\sigma_{v_n}^2)(b-a)}\right]. \tag{5.128}$$

Eq. (5.126) may then be interpreted as the stationarity result,

$$S'(v_n)h_n = 0, \quad \text{for all} \quad h_n \in \mathcal{F}([a,b]). \tag{5.129}$$

In fact, replacing v in Eq. (5.106) with h_n yields,

$$\sum_{p=1}^{n} a_p d_p = \sigma_{uh_n} + \langle u\rangle\langle h_n\rangle = \sigma_{uh_n} + \bar{u}\bar{h}_n. \tag{5.130}$$

Notice that a_p represents the "Fourier coefficients", and d_p refers to Eq. (5.123). Similarly, switching u in Eq. (5.106) to h_n gives to,

$$\sum_{p=1}^{n} c_p d_p = \sigma_{v_nh_n} + \langle v_n\rangle\langle h_n\rangle = \sigma_{v_nh_n} + \bar{v}_n\bar{h}_n, \tag{5.131}$$

where c_k is interpreted by Eq. (5.107). As a consequence, Eq. (5.126) becomes that

$$\frac{\langle h_n\rangle}{\bar{v}_n} + \frac{\sigma_{uh_n}}{\sigma_{uv_n}} - \frac{2\bar{v}_n\bar{h}_n}{\bar{u}^2+\bar{v}_n^2} - \frac{2\sigma_{v_nh_n}}{\sigma_u^2+\sigma_{v_n}^2} = 0, \tag{5.132}$$

and the Fréchet derivative of $S(v)$ is simplified as follows,

$$S'(v_n)h = S(v_n)\left[\frac{\langle h_n\rangle}{\bar{v}_n} + \frac{\sigma_{uh_n}}{\sigma_{uv_n}} - \frac{2\bar{v}_n\bar{h}_n}{\bar{u}^2+\bar{v}_n^2} - \frac{2\sigma_{v_nh_n}}{\sigma_u^2+\sigma_{v_n}^2}\right]. \tag{5.133}$$

Chapter 6

Existence and Uniqueness of Best Generalized Weber-based L^2 Approximations

In the context of optimization theory, the existence and uniqueness of the optimal solutions (i.e., minimizers or maximizers) in terms of some best approximation problems is a classical topic which has been widely considered and investigated in the field of mathematical imaging. The purpose of this chapter is primarily to examine the existence and uniqueness properties of minimizers when the generalized Weber-based distance, denoted as $D(u, v; \nu_a)$, between the reference function u and its approximation v is minimized, where $a > 0$ and ν_a represents the nonuniform intensity-based measures that accommodate generalized Weber's model of perception. Given some $a > 0$, we previously demonstrated the distance functions $D(u, v; \nu_a)$ are metrics while they do not come from the normed space. As a consequence, the theory of best approximation in normed spaces and strictly normed spaces cannot be appropriately applied in analyzing whether there exists a unique solution as $D(u, v; \nu_a)$ is minimized for any $a > 0$. But in this chapter, we shall establish an existence-uniqueness theorem of best generalized Weber-based approximation by studying the properties (e.g., compactness, convexity, etc) of a function space which contains all possible approximate functions v.

6.1 Introduction

Now, let us consider the following squared L^2 distance functions defined by the intensity-based measures ν_a,

$$D(u, v; \nu_a) = \int_X [P(u(x)) - P(v(x))]^2 \, dx, \tag{6.1}$$

where $a \geq 0$, P is the antiderivative of the continuous density function ρ_a, and for any $I_1, I_2 \in \mathbb{R}_g$,

$$\nu_a(I_1, I_2) = \int_{I_1}^{I_2} \rho_a(y) dy. \tag{6.2}$$

In what follows, we consider u as a target function, and v as the approximation which is described by a linear combination of an orthonormal basis denoted as $\{\phi_k\}_{k=1}^n$, i.e.,

$$v_n(x) = \sum_{k=1}^{n} c_k \phi_k(x). \tag{6.3}$$

Then, Eq. (6.1) becomes

$$D(c_1, \cdots, c_n) = \int_X [P(u(x)) - P(\sum_{k=1}^{n} c_k \phi_k(x))]^2 \, dx, \tag{6.4}$$

which is a multi-variable function in terms of a set of parameters $c_1, \cdots, c_n$ over C_n. As in Chapter 5, C_n, a subset of $\mathbb{R}^n$, denotes a suitably-defined feasible set of parameters: If $c = (c_1, \cdots, c_n) \in C_n$ then the linear combination $v_n(x)$ indicated above lies in the following set of positive-valued functions $\mathcal{F}(X) \subset L^2(X)$,

$$\mathcal{F}(X) = \{u \in L^2(X) \mid 0 < A \leq u(x) \leq B \;\; \text{a.e.} \;\; x \in X\}, \tag{6.5}$$

where $X \subset \mathbb{R}$, and $L^2(X)$ represents the Hilbert space. In a special case of Lebesgue measure (i.e., $a = 0$), $P(y) = y$. From Eq. (6.1), we obtain that

$$D(u, v) = \int_X [u(x) - v(x)]^2 \, dx, \tag{6.6}$$

which is the usual squared L^2 distance between the functions $u(x)$ and $v(x)$. It follows from Eq. (6.3) that the above equation becomes

$$D(c_1, \cdots, c_n) = \int_X [u(x) - \sum_{k=1}^{n} c_k \phi_k(x)]^2 \, dx. \tag{6.7}$$

Note that Eq. (6.6) is a distance function with absence of Weber's model of perception.

6.2 Existence and Uniqueness of Best Approximations

6.2.1 Existence and Uniqueness of the Optimal Solution for the Approximation of non-negative Functions via Standard L^2 Metric

It is well-known that the best approximation in Hilbert spaces is as follows [27, 28]:

Theorem 6.2.1: Let $\{\phi_k\}_{k=1}^n$ represent an orthonormal set in a Hilbert space H, and we define $Y = \text{span}\{\phi_k\}_{k=1}^n$, which is a subspace of H. Then for any $u \in H$, the best approximation of u in Y is given by the unique vector

$$y = P_Y(u) = \sum_{k=1}^{n} c_k \phi_k, \tag{6.8}$$

where

$$c_k = \langle u, \phi_k \rangle, \quad k = 1, 2, \cdots, n. \tag{6.9}$$

The c_k are called the Fourier coefficients of u with respect to the set $\{\phi_k\}_{k=1}^n$.

A detailed proof of this standard theorem can be found in many textbooks, e.g., [27]. Indeed, the key idea of the best approximation in Hilbert space is to find the point $y \in Y$ that minimizes the squared distance $||u - v||_2^2$, for all $v \in Y$, i.e.,

$$y = \arg\min_{v \in Y} ||u - v||_2^2, \tag{6.10}$$

where $||\cdot||_2$ denotes the usual L^2 norm. In other words, we have to find the optimal coefficients $c_1, c_2, \cdots, c_n$ such that the distance, i.e., Eq. (6.7), is minimized. As a consequence, the above theorem presents that there exists a unique optimal solution $y \in Y$ such that Eq. (6.7) is minimized.

Given the metric space $(\mathcal{F}(X), d_2)$, we let $S_n \subseteq \mathcal{F}(X)$ be a subset of $\mathcal{F}(X)$ (i.e., Eq. (6.5)), where the function space S_n is defined as follows,

$$S_n = \left\{ v : X \to \mathbb{R}_g \;\middle|\; v(x) = \sum_{k=1}^{n} c_k \phi_k(x) \;\; \text{for} \;\; \mathbf{c} \in C_n \right\}, \tag{6.11}$$

where $\mathbf{c} = (c_1, \cdots, c_n)$. Once again, C_n is a feasible set of parameters which makes all the approximations, i.e., $v(x)$, restricted in the set of functions $\mathcal{F}(X)$. Now, given a function

$u \in \mathcal{F}(X)$, instead of the space $Y = \text{span}\{\phi_k\}_{k=1}^n$, we are interested in searching the best approximation $v(x)$ in the above function space S_n. Mathematically, the natural question is "Does there exist a unique optimal solution $v \in S_n$ such that Eq. (6.7) is minimized?" Note that both $u(x)$ and $v(x)$ are assumed as the positive functions, i.e., $A \leq u(x), v(x) \leq B$ a.e. $x \in X$. Referring to [29], the distance of a given function $u \in \mathcal{F}(X)$ to S_n can be defined to be the number

$$\text{dist}(u, S_n) = \inf\{d_2(u, v) : v \in S_n\}. \tag{6.12}$$

Moreover, if $v_0 \in S_n$ is the minimizer such that $d_2(u, v_0) = \text{dist}(u, S_n)$, then v_0 is called a metric projection of u onto S_n. More importantly [29], in the special case of S_n is a convex body (i.e., a compact convex set with non-empty interior), we learn that S_n is called a Chebyshev set, implying that there exists a unique metric projection onto S_n, a subset of finite-dimensional Hilbert space $L^2(X)$, for any function $u \in \mathcal{F}(X)$. In what follows, we shall formally prove that the function space S_n is indeed a convex body.

At first, let us investigate the convexity of the feasible set C_n. From the definition of a convex set, for any $c = (c_1, \cdots, c_n), d = (d_1, \cdots, d_n) \in C_n \subseteq \mathbb{R}^n$, if we can prove that

$$\lambda c + (1 - \lambda)d \in C_n, \quad \lambda \in [0, 1]\,, \tag{6.13}$$

then C_n is a convex set. In fact, the only restriction making C_n considered as a feasible set of the parameters is presented in Eq. (6.5), i.e., positive-valuedness and boundedness of the approximate function $v(x)$. As a result, for any $\lambda \in [0, 1]$, we obtain that

$$\lambda A \leq \lambda[c_1\phi_1(x) + \cdots + c_n\phi_n(x)] \leq \lambda B, \tag{6.14}$$

and

$$(1 - \lambda)A \leq (1 - \lambda)[d_1\phi_1(x) + \cdots + d_n\phi_n(x)] \leq (1 - \lambda)B. \tag{6.15}$$

Adding Eq. (6.14) to Eq. (6.15) yields

$$A \leq [\lambda c_1 + (1 - \lambda)d_1]\phi_1(x) + \cdots + [\lambda c_n + (1 - \lambda)d_n]\phi_n(x) \leq B, \tag{6.16}$$

which implies that the set of parameters $(\lambda c_1 + (1 - \lambda)d_1, ..., \lambda c_n + (1 - \lambda)d_n) \in C_n$, for all $\lambda \in [0, 1]$. Hence, C_n is a convex set. Furthermore, for any $v^{(1)}, v^{(2)} \in S_n$, we have two sets of parameters $c = (c_1, \cdots, c_n), d = (d_1, \cdots, d_n) \in C_n$ such that

$$v^{(1)} = \sum_{k=1}^{n} c_k\phi_k(x), \quad v^{(2)} = \sum_{k=1}^{n} d_k\phi_k(x). \tag{6.17}$$

Then, we obtain that

$$\lambda v^{(1)} + (1-\lambda)v^{(2)} = \sum_{k=1}^{n} [\lambda c_k + (1-\lambda)d_k]\phi_k(x). \tag{6.18}$$

Since C_n is a convex set, $\lambda c + (1-\lambda)d \in C_n$. Hence, we have that $\lambda v^{(1)} + (1-\lambda)v^{(2)} \in S_n$, according to Eq. (6.11). This implies that the function space S_n is also a convex set. Likewise, we may use the same manner presented above to prove that the function space $\mathcal{F}(X)$ is convex as well. If we have two functions $u, v \in \mathcal{F}(X)$, then for any $\lambda \in [0,1]$, it follows from Eq. (6.5) that

$$\lambda A \leq \lambda u(x) \leq \lambda B, \tag{6.19}$$

and

$$(1-\lambda)A \leq (1-\lambda)v(x) \leq (1-\lambda)B, \tag{6.20}$$

for a.e. $x \in X$. As such, adding up Eqs. (6.19) and (6.20) yields

$$A \leq \lambda u(x) + (1-\lambda)v(x) \leq B, \tag{6.21}$$

which implies the convexity of $\mathcal{F}(X)$.

Next, we are going to show that for each $n \geq 1$, the set of parameters $C_n \subset \mathbb{R}^n$ is a closed and bounded set. For the boundedness of C_n, we know $v_n(x) \in S_n \subseteq \mathcal{F}(X)$. From Eqs. (6.3) and (6.5), it is clear that the range of the approximate functions is the interval $[A, B]$, i.e.,

$$A \leq v_n(x) = \sum_{k=1}^{n} c_k \phi_k(x) \leq B, \quad \text{a.e. } x \in X. \tag{6.22}$$

As before, $\{\phi_k\}_{k=1}^n$ denotes a set of orthonormal basis for the metric space. For each $n \geq 1$, we would like to find the set $C_n \subseteq \mathbb{R}^n$ such that the above inequality is satisfied. Now, we shall make some assumptions on the basis functions $\phi_k(x)$. First of all, they are assumed to be continuous on X, which implies that there exist constants m_k and M_k, $k \geq 1$, such that

$$m_k \leq \phi_k(x) \leq M_k, \tag{6.23}$$

for a.e. $x \in X$. In this particular study, with an eye to applications, we shall consider the special case of cosine functions in which case

$$\phi_1(x) = m_1 = M_1 = K, \tag{6.24}$$

and

$$m_k = -M_k = -M, \quad k \geq 2, \tag{6.25}$$

i.e.,

$$-M \leq \phi_k(x) \leq M \quad \text{for a.e. } x \in X\,, \quad k \geq 2\,. \tag{6.26}$$

As a special case of $n = 1$, it follows from Eqs. (6.22) and (6.24) that

$$A \leq c_1 K \leq B \implies \frac{A}{K} \leq c_1 \leq \frac{B}{K}, \tag{6.27}$$

implying that the feasible set C_1 is bounded by the interval $[A/K, B/K]$. When $n = 2$, Eq. (6.22) becomes that

$$A \leq c_1 K + c_2 \phi_2(x) \leq B. \tag{6.28}$$

Since the above inequality is true for a.e. $x \in X$, we obtain the following two inequalities,

$$\begin{aligned} A &\leq c_1 K + c_2 M \leq B, \\ A &\leq c_1 K - c_2 M \leq B. \end{aligned} \tag{6.29}$$

Note that the above two inequalities define the feasible set C_2, which contains the set C_1 lying on the c_1-axis. Based on the results of these two special cases, it seems that we can possibly derive bounds on the feasible set C_n by employing a recursive procedure. As a matter of fact, if we substitute the left inequality of Eq. (6.27) into Eq. (6.29), we obtain that

$$\begin{cases} A \leq A + c_2 M \leq B \\ A \leq B + c_2 M \leq B, \end{cases} \tag{6.30}$$

and

$$\begin{cases} A \leq A - c_2 M \leq B \\ A \leq B - c_2 M \leq B. \end{cases} \tag{6.31}$$

Solving both Eq. (6.30) and Eq. (6.31) yields the inequality of c_2 as follows,

$$\frac{A - B}{M} \leq c_2 \leq \frac{B - A}{M}. \tag{6.32}$$

Now, let us consider the inequality for general $n \geq 1$. From Eqs. (6.22) and (6.24), we have that

$$A \leq c_1 K + c_2 \phi_2(x) + \cdots + c_n \phi_n(x) \leq B. \tag{6.33}$$

Then, for $n + 1$,

$$A \leq [c_1 K + c_2 \phi_2(x) + \cdots + c_n \phi_n(x)] + c_{n+1} \phi_{n+1}(x) \leq B. \tag{6.34}$$

Similarly, if we employ the two extremes in Eq. (6.33) for the expression in brackets as we did in earlier for $n = 1$ and $n = 2$ to obtain the inequalities,

$$\frac{A - B}{M} \leq c_{n+1} \leq \frac{B - A}{M}, \quad n \geq 1. \tag{6.35}$$

By induction, the above inequality holds for all c_n, where $n \geq 2$. For $n = 1$, c_1 is bounded as indicated in Eq. (6.27). Hence, we have now proved that the set of parameters C_n is bounded.

In order to establish that C_n is closed, let us define a convergent sequence $\{c_{jk}\} \subset C_n$ such that

$$\lim_{j\to\infty} c_{jk} = d_k, \tag{6.36}$$

for all $k = 1, \cdots, n$. Since $\{c_{jk}\} \subset C_n$, we have that

$$A \leq \sum_{k=1}^{n} c_{jk}\phi_k(x) \leq B, \quad \text{a.e. } x \in X. \tag{6.37}$$

Now, if we let

$$f_j(x) = \sum_{k=1}^{n} c_{jk}\phi_k(x), \tag{6.38}$$

and

$$f(x) = \sum_{k=1}^{n} d_k\phi_k(x), \tag{6.39}$$

then it is straightforward to see that

$$\lim_{j\to\infty}[f_j(x) - f(x)] = \lim_{j\to\infty}\sum_{k=1}^{n}(c_{jk} - d_k)\phi_k(x) = 0, \quad \text{a.e. } x \in X. \tag{6.40}$$

Therefore, as $n \to \infty$, Eq. (6.37) becomes

$$A \leq \sum_{k=1}^{n} d_k\phi_k(x) \leq B, \quad \text{a.e. } x \in X. \tag{6.41}$$

This suggests that $d_k \in C_n$ for all $k = 1, \cdots, n$, implying that $C_n \subset \mathbb{R}^n$ is a closed set. From Bolzano-Weierstrass Theorem [30], we shall conclude that the set of parameters $C_n \subset \mathbb{R}^n$ is a compact set since C_n is closed and bounded.

From Proposition 1.9 in [26], we learn that any nonempty convex set $C_n \subset \mathbb{R}^n$ has a

non-empty relative interior. Moreover, it is clear that $\dim C_n = n$, which implies that the convex set $C_n \subset \mathbb{R}^n$ is full dimension. As such, the convex set C_n has a nonempty interior [26]. We now have proved the set of parameters C_n is a convex compact set with nonempty interior, meaning that C_n is a convex body. As we know, all the basis functions $\phi_k(x)$, $k = 1, \cdots, n$, are continuous. Hence, the function space S_n is also a convex body in that S_n is homeomorphic to the set C_n. Overall, we have presented that the set S_n described by Eq. (6.11) is a Chebyshev set, which completes the proof of existence and uniqueness of the optimal solution in S_n when the usual L^2 distance function is minimized.

6.2.2 Existence and Uniqueness of the Optimal Solution for the Approximation of non-negative Functions via Weber-based L^2 Metrics

Now, we consider the following Weberized distance functions/metrics on the space $\mathcal{F}(X)$:

- For $0 < a < 1$:

$$D_a(c_1, \cdots, c_n) = \int_X [u(x)^{-a+1} - (\sum_{k=1}^{n} c_k \phi_k(x))^{-a+1}]^2 dx. \tag{6.42}$$

- For $a = 1$:

$$D_1(c_1, \cdots, c_n) = \int_X [\log u(x) - \log(\sum_{k=1}^{n} c_k \phi_k(x))]^2 dx. \tag{6.43}$$

As usual, the metric $D_1(u, v_n)$ presented above represents the standard Weberized L^2 distance function, which corresponds to the standard Weber's model of perception; and both $D_a(c_1, \cdots, c_n)$ and $D_1(c_1, \cdots, c_n)$ are continuous on the set of coefficients C_n. If we recall some results obtained in the last section, i.e., Section 6.2.1, $C_n \subset \mathbb{R}^n$ is a compact set, which consequently implies that there must exist at least one solution when we wish to find the best approximation of the reference function $u(x) \in \mathcal{F}(X)$ by minimizing the distance functions in Eqs. (6.42) and (6.43).

For the above two metrics, let us consider the following functions:

$$U_1(x) = \log(u(x)), \quad U_2(x) = u(x)^{-a+1}, \tag{6.44}$$

and

$$V_{1,n}(x) = \log(\sum_{k=1}^{n} c_k \phi_k(x)), \quad V_{2,n}(x) = (\sum_{k=1}^{n} c_k \phi_k(x))^{-a+1}, \tag{6.45}$$

where $x \in X$. Since $u(x)$ and $v_n(x) = \sum_{k=1}^{n} c_k \phi_k(x)$ are in the function space $\mathcal{F}(X) \subset L^2(X)$, it is clear that $U_1(x), U_2(x), V_{1,n}(x), V_{2,n}(x) \in \mathcal{F}(X) \subset L^2(X)$ as well. As such, the best approximation problems investigated on minimizing the two metrics, i.e., Eqs. (6.42) and (6.43) are still centered in the $L^2(X)$ space, i.e., the Hilbert space. This is actually why we sometimes call the two metrics in Eq. (6.42) and Eq. (6.43) modified L^2 distance functions. Based on the fact that each $u(x) \in \mathcal{F}(X)$ has at least one best approximation in the set S_n (or C_n), S_n (or C_n) is now called a proximinal set [28]. Furthermore, since S_n is also a convex set as we proved in the last section, we may employ the following theorem [28] to show the uniqueness of best approximations via minimizing the distance functions described in Eqs. (6.42) and (6.43).

Theorem 6.2.2: Let K be a convex subset of a Hilbert space H. Then each $u \in H$ has at most one best approximation of u in K. In particular, every convex proximinal set is Chebyshev.

Apparently, S_n is a Chebyshev set, which is also a subset of the function space $\mathcal{F}(X)$. This completes the proof of the uniqueness of the approximation $v_n(x)$ in S_n while the two metrics, i.e., Eqs. (6.42) and (6.43) are minimized.

In what follows, we shall consider a special case of the best affine logarithmic L^2 approximation in the Hilbert space $H = L^2([0,1])$, in which case

$$u(x) = x^2 + 1, \quad v(x) = cx + d. \tag{6.46}$$

From Eq. (6.43), the distance function that is relevant to the best affine logarithmic L^2 approximation is then described as follows,

$$D_1(c,d) = \int_0^1 [\log(x^2+1) - \log(cx+d)]^2 dx. \tag{6.47}$$

For each pair $c, d \in \mathbb{R}$, we are approximating $u(x)$ using $v_2(x) \in S_2 \subset \mathcal{F}([0,1])$, where

$$\mathcal{F}([0,1]) = \{f \in L^2([0,1]) \mid 0 < A \leq f(x) \leq B \ \text{ a.e. } \ x \in [0,1]\}, \tag{6.48}$$

and S_2 refers to Eq. (6.11) for $n = 2$. Of course, the reference function $u(x)$ has to be contained in the space $\mathcal{F}([0,1])$ as well. By imposing the stationarity conditions, we obtain the gradient of $D_1(c,d)$, $\nabla D_1(c,d)$, as follows,

$$\begin{aligned} \frac{\partial D_1(c,d)}{\partial c} &= \int_0^1 [\log(x^2+1) - \log(cx+d)] \frac{x}{cx+d} dx, \\ \frac{\partial D_1(c,d)}{\partial d} &= \int_0^1 \Big[\log(x^2+1) - \log(cx+d)\Big] \frac{1}{cx+d} dx, \end{aligned} \tag{6.49}$$

which yields a nonlinear system of equations with respect to the parameters c and d. Intuitively, one may first try employing the steepest-descent approach to locate the minima of $D_1(c,d)$. Starting at a suitable "seed" point (c_0, d_0), let us consider the following elementary procedure to travel in the direction of steepest descent,

$$(c_{n+1}, d_{n+1}) = (c_n, d_n) - h_n \nabla D_1(c_n, d_n), \tag{6.50}$$

where the constant h_n denotes the step size for each iteration. At each (c_n, d_n), the integral in Eq. (6.50), i.e. the approximation error, shall be evaluated numerically.

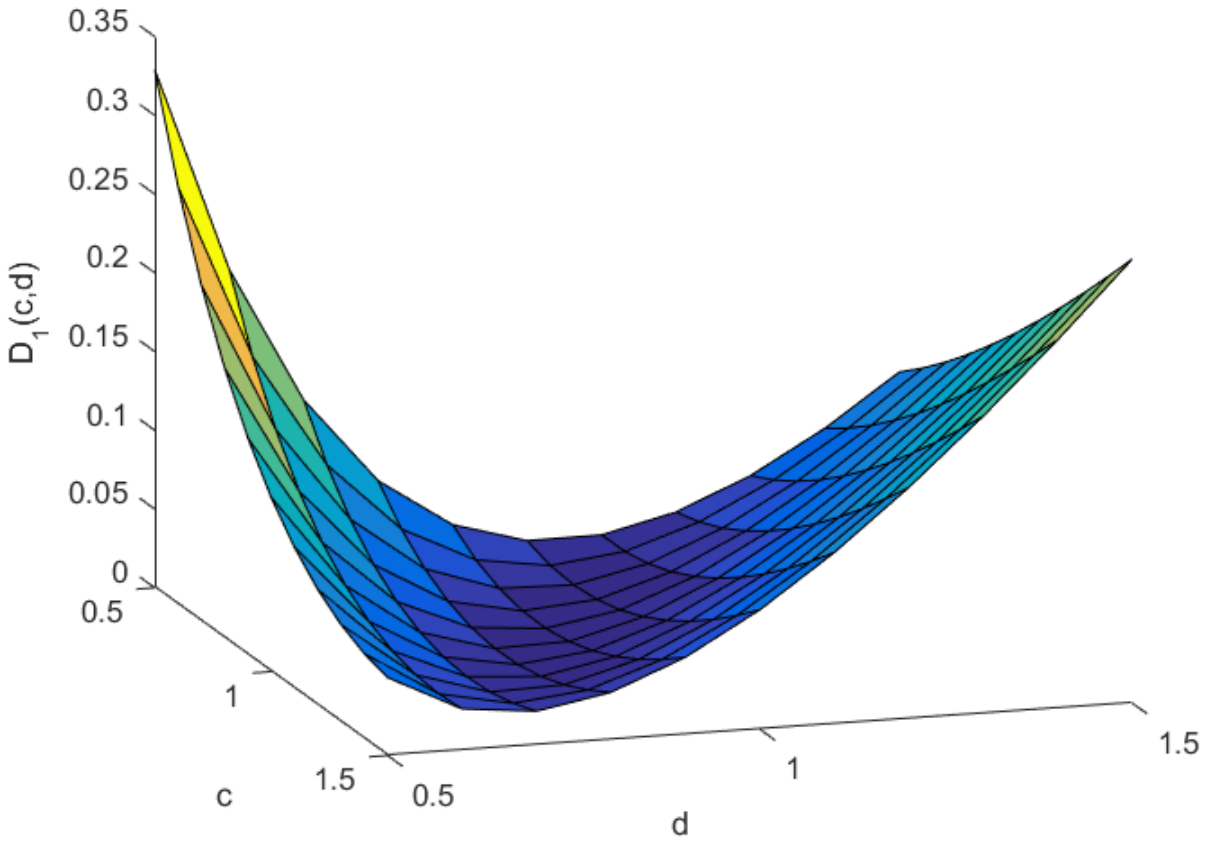


Figure 6.1: Surface plot of the affine logarithmic L^2-type distance function $D_1(c,d)$

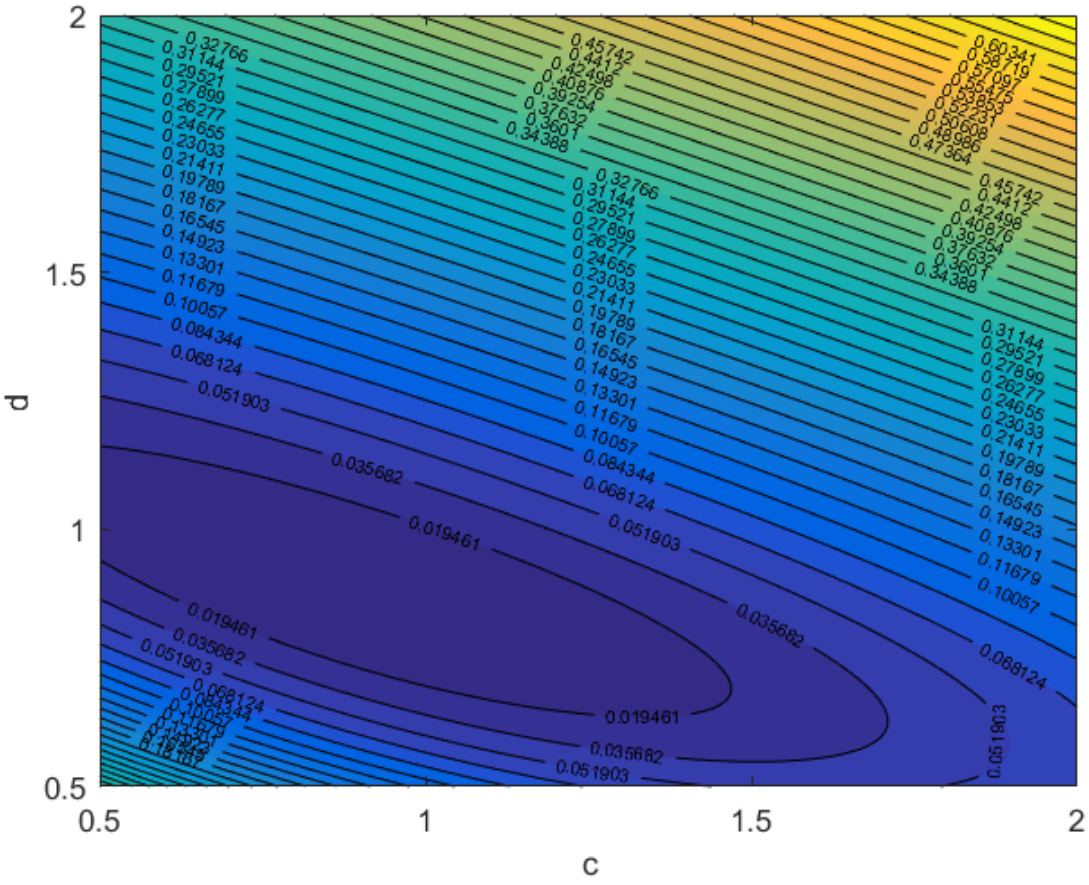


Figure 6.2: Contour plot of the affine logarithmic L^2-based metric $D_1(c, d)$

Using MATLAB, the numerical solution to c and d are found to be, to four decimals,

$$\begin{cases} c = 0.9056 \\ \\ d = 0.8757, \end{cases} \tag{6.51}$$

which implies the affine logarithmic L^2 approximation of $u(x) = x^2 + 1$ is

$$v_2(x) = 0.9056x + 0.8757. \tag{6.52}$$

To visualize the distance presented in Eq. (6.47), we make a 3D plot (Figure 6.1) of the affine logarithmic distance with x-axis, y-axis and z-axis represents c, d and $D_1(c, d)$,

respectively. Notice that c and d are both locally bounded by the interval $[0.5, 1.5]$. In practice, this interval should also correspond to the boundedness of the set C_2, which has been fruitfully discussed in the last section. Moreover, we also make a contour plot (Figure 6.2) of the distance function $D_1(c, d)$. Note the number on each contour line represents the value of $D_1(c, d)$. If we look carefully at Figures 6.1 and 6.2, we shall find that the darkest blue region appears at around the point $(c, d) = (0.9, 0.9)$, which roughly agrees with the numerical results, i.e., Eq. (6.51), yielded by steepest-descent algorithm. In addition, these two plots also illustrate that there is at most one (locally) optimal pair of coefficients (c, d) for the best affine logarithmic L^2 approximation, which is consistent with previously discussed existence and uniqueness theorem of the minimizer in (generalized) Weberized L^2 approximation.

Gauss-Newton Algorithm and its Application on the Best Affine Logarithmic L^2 Approximation

For some discrete cases (e.g., digital signals or images), Eq. (6.43) can be modified as follows,

$$D_1(c_1, \cdots, c_n) = \sum_{i=1}^{m} [\log(u(x_i)) - \log(\sum_{k=1}^{n} c_k \phi_k(x_i))]^2, \tag{6.53}$$

where $m \geq n$. Now, our best approximation problems in the usual l^2 space are to search a "best" set of parameters $c_1, \cdots, c_n$ minimizing $D_1(c_1, \cdots, c_n) = \sum_{i=1}^{m} r_i^2$, where

$$r_i = \log(u(x_i)) - \log(\sum_{k=1}^{n} c_k \phi_k(x_i)), \quad i = 1, \cdots, m. \tag{6.54}$$

In the theory of optimization, the above best approximation problem is considered as a nonlinear least squares minimization problem. Indeed, there is an iterative algorithm, so-called the Gauss-Newton algorithm, which is designed to solve the nonlinear least squares problem by applying the least-squares method [33]. Typically, it is seen as an approximation of Newton's method for finding the minimum of a function.

In this nonlinear least squares minimization problem, we aim to determine the optimal coefficients $\mathbf{c} = (c_1, \cdots, c_n)$ that minimize

$$\Delta(\mathbf{c}) = \|r(x)\|^2 = \sum_{i=1}^{m} r_i^2(\mathbf{c}), \tag{6.55}$$

where $r : \mathbb{R}^n \to \mathbb{R}^m$. Then, it follows from Eq. (6.55) that the gradient g_k for each parameter c_k is described as follows,

$$g_k = 2\sum_{i=1}^{m} r_i \frac{\partial r_i}{\partial c_k}, \tag{6.56}$$

and the Hessian for each c_k is given by

$$H_{kl} = 2\sum_{i=1}^{m} \left(\frac{\partial r_i}{\partial c_k}\frac{\partial r_i}{\partial c_l} + r_i \frac{\partial^2 r_i}{\partial c_k \partial c_l} \right), \tag{6.57}$$

where $k = 1, \cdots, n$ and $l = 1, \cdots, n$. Note that the second term in Eq. (6.57) is (nearly) zero when we are solving the linear (approximately linear) least squares well-posed problem. In essence, the Gauss-Newton method approximates the nonlinear least squares problem by some kind of the first order linearization. As such, the Hessian in Eq.(6.56) can be approximated as follows,

$$H_{kl} \approx 2\sum_{i=1}^{m} \frac{\partial r_i}{\partial c_k}\frac{\partial r_i}{\partial c_l}, \tag{6.58}$$

when we use Gauss-Newton algorithm to solve our nonlinear least squares minimization problem. Once again, the reason why we could ignore the second term in Eq. (6.56) is because we employ the first order linear approximation of $r(x)$ in Gauss-Newton algorithm, which leads to $\dfrac{\partial^2 r_i}{\partial c_k \partial c_l}$ is fairly small in every numerical iteration.

In matrix notation, the gradient and the approximated Hessian can be expressed in the following,

$$\begin{aligned} \mathbf{g} &= 2\mathbf{J}^T\mathbf{r}, \\ \mathbf{H} &\approx 2\mathbf{J}^T\mathbf{J}, \end{aligned} \tag{6.59}$$

where

$$\mathbf{r} = \begin{bmatrix} r_1 \\ r_2 \\ \vdots \\ r_m \end{bmatrix}, \quad \mathbf{J} = \begin{bmatrix} \dfrac{\partial r_1}{\partial c_1} & \dfrac{\partial r_1}{\partial c_2} & \cdots & \dfrac{\partial r_1}{\partial c_n} \\ \dfrac{\partial r_2}{\partial c_1} & \dfrac{\partial r_2}{\partial c_2} & \cdots & \dfrac{\partial r_2}{\partial c_n} \\ \vdots & \vdots & \ddots & \vdots \\ \dfrac{\partial r_m}{\partial c_1} & \dfrac{\partial r_m}{\partial c_2} & \cdots & \dfrac{\partial r_m}{\partial c_n} \end{bmatrix}. \tag{6.60}$$

Based on the iteration results for stationary points derived by Newton's method [32], the iteration function of the Gauss-Newton algorithm is described as follows,

$$\begin{aligned}\mathbf{c}^{(s+1)} &= \mathbf{c}^{(s)} - (\mathbf{J}^T\mathbf{J})^{-1}\mathbf{J}^T\mathbf{r} \\ &= \mathbf{c}^{(s)} - \mathbf{J}^\dagger\mathbf{r},\end{aligned} \tag{6.61}$$

where

$$\mathbf{J}^\dagger = (\mathbf{J}^T\mathbf{J})^{-1}\mathbf{J}^T, \tag{6.62}$$

named as generalized inverse of the matrix $\mathbf{J}$. Note that the term $-\mathbf{J}^\dagger\mathbf{r}$ is also called the Gauss-Newton direction that leads the iterations to the local minimum.

In what follows, let us recall the best affine logarithmic L^2 approximation, which is to minimize the distance function $D_1(c,d)$ in Eq. (6.47), i.e.,

$$D_1(c,d) = ||F(c,d)||_2^2, \tag{6.63}$$

where

$$F(c,d) = \log(cx+d) - \log(x^2+1). \tag{6.64}$$

If $F'(c,d)$ represents the derivative of $F(c,d)$ at the current estimate $c,d \in \mathbb{R}$, then the first order approximation of F at $(c+\delta c, d+\delta d)$ is given by

$$F(c+\delta c, d+\delta d) = F(c,d) + F'(c,d)\begin{bmatrix}\delta c \\ \delta d\end{bmatrix}. \tag{6.65}$$

As a result, the first order approximation for our problem at the current estimate (c,d) is obtained by linearizing the function $D_1(c,d)$, i.e.,

$$\min \quad \int_0^1 \left[\log(x^2+1) - \log(cx+d) + \left(\frac{x\delta c}{cx+d} + \frac{\delta d}{cx+d}\right)\right]^2 dx. \tag{6.66}$$

It follows from Eq. (6.64) that the Jacobian of $F(c,d)$, simply denoted as V, is given by

$$F'(c,d) = V = [v_1(x), v_2(x)] = \left[\frac{x}{cx+d}, \frac{1}{cx+d}\right] \tag{6.67}$$

As introduced before, the Gauss-Newton direction d_{GN} at the current estimate $c,d \in \mathbb{R}$ gives to

$$d_{GN} = V^\dagger(-F(c,d)), \tag{6.68}$$

where

$$V^\dagger = (V^*V)^{-1}V^*. \tag{6.69}$$

Note that V^* denotes the adjoint of V, so that

$$V^*V = \begin{bmatrix} ||v_1(x)||^2 & \langle v_1(x), v_2(x)\rangle \\ \langle v_1(x), v_2(x)\rangle d & ||v_2(x)||^2 \end{bmatrix}. \tag{6.70}$$

From the optimality condition, i.e., $V^TV = 0$, which can be viewed as a stopping criterion, we implement Gauss-Newton algorithm for this special problem and obtain the same result as Eq. (6.51) just in 9 iterations. Comparing with steepest descent method, Gauss-Newton algorithm dramatically improves the efficiency of optimization process for solving the best affine logarithmic L^2 approximation, since it can achieve quadratic rate of (local) convergence under certain regularity conditions, e.g., the initial iteration is reasonably close to the solution, [34]. Generally, the efficiency of Gauss-Newton algorithm could be extended to higher dimensional ($n > 2$) best approximation problems in terms of both signals and images.

Chapter 7

Concluding Remarks

In this chapter, we first present a summary of the contributions of this thesis. For extensions of this thesis, we also discuss some possible future work in terms of both mathematical imaging analysis and the applications of image processing.

7.1 Summary of Contributions

The motivation of this thesis stems from realizing the challenges of a big gap between image analysis and human visual system. As such, this thesis devotes to constructing the metrics which can predict human visual perception of image distortions. In practice, we aim to effectively apply these metrics in best approximation problems to improve perceptual quality of compressed or approximate images.

Indeed, the first contribution of this thesis was to generalize Weber's model of perception and derive a general form of "Weberized" distance functions. Secondly, we showed how to construct generalized Weber-based metrics using nonuniform intensity-based measures which accommodate a generalized Weber's model of perception in terms of an "equal area" rule. From mathematical perspective, the existence-uniqueness theorem of the density functions associated with the intensity-based measures was proved via a well-known functional equation — Abel's equation, which controls the iteration dynamics of the embedded function. We also derived the leading asymptotic behaviour of these density functions and then computed their asymptotic expansions. Furthermore, we analyzed the Fréchet differentiability of generalized Weber-based metrics and determined the Fréchet derivatives of these distance functions. We then employed the Fréchet derivatives of generalized Weber-based metrics ($0 < a \leq 1$) to examine the relation between the stationary points

and minimizers of these metrics. We saw that the stationary points of the generalized Weber-based metrics (i.e., $0 < a < 1/2$ and $a = 1$) are indeed the minimizers of these metrics if the lower and upper bounds of greyscale range satisfy some particular conditions (i.e., Eqs. (5.93) and (5.102)). In Chapter 6, the existence-uniqueness theorem of best generalized Weber-based approximation stated that every "convex body" of finite dimensional Hilbert space is proximinal and Chebyshev.

7.2 Future Directions

Weber's model of perception and its generalization allow greater deviations at higher intensity values and lower deviations at lower intensity values. Motivated by this property, one may expect to obtain more variance at lower intensities and lesser variance at higher intensities when the "opposite effect" of Weber's (generalized) model of perception incorporates in the construction of distance metrics. This "opposite effect" leads to a kind of "anti-Weber" model of perception in which the component a in Eq. (2.3) is negative. Consequently, we are able to establish another class of metrics which accommodate anti-Weber's model of perception. For the purpose of image processing applications, it would be interesting to find a hybrid image approximation model which can adaptively yield Weberized and anti-Weberized approximations with respect to lower and higher intensity regions of an image. In the future, one may also carry out a number of objective and subjective experiments to test the performance of generalized Weber-based metrics, which could provide some help with image quality assessment. Naturally, these experiments should be designed for indicating the effect that generalized Weber-based metrics are in accordance with human visual system.

In fact, our intensity-based measure approach does not have to be associated with Weber or anti-Weber models of perception. It represents a mathematical method that can be used to construct arbitrary range-based measures that may be used in function approximation. As such, it may have applications that lie well beyond the scope of mathematical imaging.

References

[1] I.A. Kowalik-Urbaniak, D. La Torre, E.R. Vrscay and Z. Wang. Some "Weberized" L^2-based methods of signal/image approximation. In *Image Analysis and Recognition*, ICIAR 2014, Part I, LNCS 8814, pp. 20-29, 2014.

[2] I.A. Kowalik-Urbaniak. *The quest for diagnostically lossless medical image compression using objective image quality measures.* PhD thesis, University of Waterloo, Waterloo, ON, Canada, 2015.

[3] D. Li, D. La Torre and E.R. Vrscay. The use of intensity-based measures to produce image function metrics which accommodate Weber's models of perception. In *Image Analysis and Recognition*, ICIAR 2018, LNCS 10882, 326-335, 2018.

[4] B. Forte and E.R. Vrscay. Solving the inverse problem for function and image approximation using iterated function systems. Dynamics of Continuous, Discrete and Impulsive Systems 1, 177-231, 1995.

[5] I. Mishra, D. Sharma and S. Jain. Compressed Sensing: Progresses and Challenges. *International Journal of Scientific Engineering and Applied Science (IJSEAS)*, ISSN: 2395-3470, 2015.

[6] Shalev-Shwartz and A. Tewari. Stochastic methods for ℓ_1 regularized loss minimization. In *Proceedings of the 26th International Conference on Machine Learning (ICML)*, pages 929936, Montreal, Canada, 2009.

[7] J. Shen. On the foundations of vision modeling I. Weber's law and Weberized TV restoration. Physica D 175 241-251, 2003.

[8] J. Shen and Y.M. Jung. Weberized Mumford-Shah model with Bose-Einstein photon noise. Appl. Math. Optim. 53 331-358, 2006.

[9] S. Hecht. The visual discrimination of intensity and the weber-fechner law. *Journal of General Physiology*, 7(2):235267, 1924.

[10] T. N. Cornsweet, H. M. Pinsker. Luminance Discrimination of Brief Flashes under Various Conditions of Adaptation. *Journal of Physiology* 176:294310, 1965.

[11] F. Kingdom, P. Whittle. Contrast discrimination at high contrasts reveals the influence of local light adaptation on contrast processing. Vis. Res., vol. 36, no. 6, pp. 817-829, 1996.

[12] A. R. Halpern, C. J. Darwin. Duration discrimination in a series of rhythmic events. Perception and Psychophysics, 31, 8689, 1982.

[13] D. Li, D. La Torre and E.R. Vrscay. Existence, uniqueness and asymptotic behaviour of intensity-based measures which conform to Weber's models of perception. *Image Analysis and Recognition*, ICIAR 2019, LNCS 11662, 297-308, 2019.

[14] W. Forg-Rob, D. Gronau, C. Mira, N. Neizer and Gy. Targonski. *Iteration theory : proceedings of the European Conference*, Batschuns, Austria, 1992.

[15] R. Tambs Lyche. Sur l'équation d'Abel. Fund. Math. 5, 331-333, 1924.

[16] G. Belitskii and Y. Lyubich. The Abel equation and total solvability of linear functional equations. Studia Mathematica 127 (1), 81-97, 1998.

[17] D. Brunet, E.R. Vrscay and Z. Wang. Structural similarity-based approximation of signals and images using orthogonal bases. In *Image Analysis and Recognition*, ICIAR 2010, LNCS 6111, Springer, Heidelberg, pp. 11-22.

[18] E. H. Weber. *De pulsu, resorptione, audita et tactu. Annotationes anatomicae et physiologicae*. Koehler, Leipzig, 1834.

[19] G. T. Fechner. Über ein wichtiges psychophysiches Grundgesetz und dessen Beziehung zur Schäzung der Sterngrössen. Abh. Ges. Wissensch. Math.-Phys., K1, 4, 1858.

[20] N. Ahmed. How I came up with the discrete cosine transform. Digital Signal Processing Vol. 1, No. 1, pp. 4-9, 1991.

[21] Z. Wang and A. C. Bovik. Mean squared error: Love it or leave it? A new look at signal fidelity measures. *IEEE Signal Process. Mag.*, vol. 26, no. 1, pp. 98-117, 2009.

[22] T. Richter and K. J. Kim. A MS-SSIM optimal JPEG 2000 encoder. In *Data Compression Conference*, pages 401410, Snowbird, Utah, March 2009.

[23] S. S. Channappayya, A. C. Bovik, and R. W. Heath. Rate bounds on ssim index of quantized images. *IEEE Trans. Image Processing*, 17(6):857872, September 2008.

[24] P. Zhao, Y. Liu, J. Liu, R. Yao and S. Ci. SSIM-based cross-layer optimized video streaming over LTE downlink. *Proc. IEEE Globecom*, pp. 1394-1399, 2014.

[25] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli. Image quality assessment: From error visibility to structural similarity. *IEEE Trans. Image Processing*, 13(4):600612, April 2004.

[26] H. Tuy. *Convex Analysis and Global Optimization. Nonconvex Optimization and Its Applications*. Kluwer Academic Publishers, 1998.

[27] E.R. Vrscay. Best Approximation in Hilbert Space. AMATH 731 Course Notes, Department of Applied Mathematics, University of Waterloo, 2018.

[28] F. Deutsch. *Best Approximation in Inner Product Spaces*. CMS Books in Mathematics, Vol. 7, Springer-Verlag, New York, 2001.

[29] V. Balestro, H. Martini, R. Teixeira. Convex analysis in normed spaces and metric projections onto convex bodies. Functional Analysis, 2019.

[30] G. Oman. A short proof of the bolzano-weierstrass theorem. The College Mathematics Journal, 2017.

[31] A. Nedić. Fundamental Concepts in Convex Optimization. Optimization I, Lecture Notes, 2008.

[32] P. Deuflhard. A Short History of Newtons Method. Documenta Mathematica Extra Volume ISMP 25-30, 2012.

[33] J.V. Burke and M.C. Ferris. A Gauss-Newton method for convex composite optimization. Mathematical Programming, 71, 179- 194, 1995.

[34] R. W. Siregar, T. Tulus and M. Ramli. Analysis Local Convergence of Gauss-Newton Method. In *IOP Conference Series: Materials Science and Engineering*, 300, 2018.

[35] O. Kosheleva, V. Kreinovich, and H. T. Nguyen. On the Optimal Choice of Quality Metric in Image Compression. In *Proceedings of the IEEE Southwest Symposium on Image Analysis and Interpretation*, pp. 116120, 2002.

[36] Q. Huynh-Thu and M. Ghanbari. Scope of Validity of PSNR in Image/Video Quality Assessment. Electronics Letters. Vol. 44, Issue 13, pp. 800801, 2008.

[37] S. Hu, L. Jin, H. Wang, Y. Zhang, S. Kwong and C.-C.J.Kuo,. Compressed image quality metric based on perceptually weighted distortion. *IEEE Transactions on Image Processing*, 2015.

[38] N. Ye, M. Perez-Ortiz and R.K.Mantiuk. Visibility Metric for Visually Lossless Image Compression. In *Proceedings of the IEEE Picture Coding Symposium*, 2019.

[39] X. Zhao, R. Wu, Z. Zhou, and W. Wu. A new metric for measuring image-based 3D reconstruction. In *Proceedings of International Conference on Pattern Recognition*, Tsukuba Science City, Japan, 2012.

[40] N. Abel. Untersuchung der Functionen zweier unabhängig veränderlichen Größen x und y, wie $f(x, y)$, welche die Eigenschaft haben, daß$f(z, f(x, y))$ eine symmetrische Function von z, x und y ist. *J. Reine Angew. Math*, 11-15, 1826.

[41] B. Girod. What's wrong with mean squared error?. In *Digital Images and Human Vision*, A.B. Watson (Ed.), MIT Press, Cambridge MA, 1993.

[42] C. Wachinger and N. Navab. Simultaneous Registration of Multiple Images: Similarity Metrics and Efficient Optimization. In *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 35, no. 5, pp. 1221-1233, 2013.

[43] W. Forst and D. Hoffmann. *Optimization—Theory and Practice.* Springer, 2010.

[44] J. Jahn. *Introduction to the Theory of Nonlinear Optimization.* Springer-Verlag, Berlin, 1996.