"Range-based" function approximation using moments in range space and generalized Weber's model of perception)

Dongchang Li[1] Davide La Torre[2] Edward R. Vrscay[1]

Department of Applied Mathematics, Faculty of Mathematics,
University of Waterloo, Waterloo, ON, Canada N2L 3G1
d2351li@edu.uwaterloo.ca, ervrscay@uwaterloo.ca

Department of Economics, Business and Statistics,
University of Milan, Milan Italy
SKEMA Business School, 06902 Sophia Antipolis, France
davide.latorre@unimi.it

University of Waterloo

AMMCS 2019
Applied Analysis and Inverse Problems Special Session
August 20, 2019

Primary motivation: perceptual image quality measures

This study represents ongoing work on the development of nonstandard distance functions/metrics between (image) functions. One of the principal motivations for this work is the important problem of objectively measuring **perceptual image quality**.

Example: Traditional best L^2 approximation of a function u in terms of a set of N basis functions.

$$\text{Minimize} \quad \int_a^b \left[u(x) - \sum_{k=1}^{N} c_k \phi_k(x) \right]^2 dx \,. \qquad (1)$$

For reasons that may become clearer later in this talk, we may wish to solve the following approximation problem,

$$\text{Minimize} \quad \int_a^b \left[\sqrt{u(x)} - \sqrt{\sum_{k=1}^{N} c_k \phi_k(x)} \right]^2 dx \,. \qquad (2)$$

Primary motivation: perceptual image quality measures

It is well known that L^2-based metrics perform poorly in terms of perceptual quality. Much work has been done to develop better image quality measures, most notably, the **Structural Similarity** (SSIM) measure,

> *Z. Wang, A.C. Bovik, H.R. Sheikh and E.P. Simoncelli, Image quality assessment: From error visibility to structural similarity, IEEE Trans. Image Proc.* **13** *(4), 600-612 (2004).*

SSIM generally performs better for two principal reasons:

- One of its three (multiplicative) components is the **correlation** between two image blocks/patches. As such, correlation – extremely important in visual perception – plays a more direct role in SSIM than in the L^2 distance.
- The algebraic forms of the terms composing SSIM – ratios – are designed to accommodate Weber's law/model of perception, e.g.,

$$S_1(\bar{\mathbf{x}}, \bar{\mathbf{y}}) = \frac{\bar{\mathbf{x}}\bar{\mathbf{y}}}{\bar{\mathbf{x}}^2 + \bar{\mathbf{y}}^2} = \frac{1 + \bar{\mathbf{y}}/\bar{\mathbf{x}}}{1 + (\bar{\mathbf{y}}/\bar{\mathbf{x}})^2} \,. \tag{3}$$

The SSIM index is a **similarity measure**, i.e., the SSIM between two image patches $\mathbf{x}$ and $\mathbf{y}$ behaves as follows,

$$-1 \leq S(\mathbf{x}, \mathbf{y}) \leq 1\,. \tag{4}$$

This suggests that the function

$$T(\mathbf{x}, \mathbf{y}) = 1 - S(\mathbf{x}, \mathbf{y}) \implies 0 \leq T(\mathbf{x}, \mathbf{y}) \leq 2\,, \tag{5}$$

could be related to a distance function. In fact, in the case of zero-mean patches, i.e., $\bar{\mathbf{x}} = \bar{\mathbf{y}} = 0$,

$$T(\mathbf{x}, \mathbf{y}) = \frac{\|\mathbf{x} - \mathbf{y}\|_2^2}{\|\mathbf{x}\|_2^2 + \|\mathbf{y}\|_2^2}\,. \tag{6}$$

Therefore, $\sqrt{T(\mathbf{x}, \mathbf{y})}$ is a **normalized** metric. It may also be viewed as an **intensity-weighted** metric.

Once again, this accommodates Weber's model of perception.

In this work, however, we are not concerned with SSIM.

Other efforts to incorporate Weber's model of perception in image processing methods

There have been other efforts to incorporate Weber's model of perception in image processing, e.g.,

> *J. Shen, On the foundations of vision modeling I. Weber's law and Weberized TV restoration, Physica D* **175***, 241-251 (2003).*

The basic idea is to "Weberize" the method by dividing by the total variation (TV) $\|\nabla u\|$ by u to produce a modified TV, $\dfrac{\|\nabla u\|}{u}$, which assigns lower/higher weight in regions of higher/lower image intensity u. The TV is now **intensity-dependent**.

The term "Weberize" comes from Weber's law/model of perception:

$$\frac{\Delta I}{I} = C\,, \tag{7}$$

where

- $I > 0$: greyscale background intensity,
- ΔI: minimum change in intensity perceived by human visual system (HVS),
- C: constant, or at least roughly constant, over a significant range of intensities $I > 0$.

In other words, HVS is less/more sensitive to given change in intensity $\Delta I > 0$ in regions of an image at which local intensity $I(x)$ is high/low.

Other efforts to incorporate Weber's model of perception in image processing methods

A "Weberized" method, therefore, should tolerate greater/lesser differences between two functions u and v over regions in which they assume higher/lower intensity values.

This idea motivated our work presented at the ICIAR 2014 Conference (Portugal):

> *I.A. Kowalik-Urbaniak, D. La Torre, E.R. Vrscay and Z. Wang, Some "Weberized" L^2-based methods of signal/image approximation,* Image Analysis and Recognition, *ICIAR 14, LNCS* **8814**, *20-29 (2014).*

There, we considered the "Weberization" of the L^2 distance between two functions u and v – or any metric involving an integration over some power of $|u(x) - v(x)|$ – by introducing an **intensity-based weight function** into the integration, as outlined below.

Other efforts to incorporate Weber's model of perception in image processing methods

One way to "Weberize" the L^2 distance

Consider the usual L^2 distance between two image functions,

$$d_2(u, v) = \left[\int_a^b [u(x) - v(x)]^2 \, dx \right]^{1/2}. \tag{8}$$

If we consider $u(x)$ to be a reference function and $v(x)$ its approximation, introduce $\frac{1}{u(x)^2}$ as intensity-dependent weighting function,

$$\begin{aligned} \Delta(u, v) &= \left[\int_a^b \frac{1}{u(x)^2} [u(x) - v(x)]^2 \, dx \right]^{1/2} \\ &= \left[\int_a^b \left[1 - \frac{v(x)}{u(x)} \right]^2 dx \right]^{1/2}. \end{aligned} \tag{9}$$

If we let

$$v_N(x) = \sum_{k=1}^{N} c_k \phi_k(x) \tag{10}$$

be an approximation to $u(x)$, then minimization of $\Delta(u, v_N)$ yields linear system of equations in c_k, $1 \le k \le N$.

Other efforts to incorporate Weber's model of perception in image processing methods

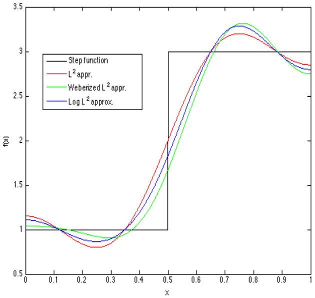
Another way to "Weberize" the L^2 distance: "Logarithmic L^2 distance"

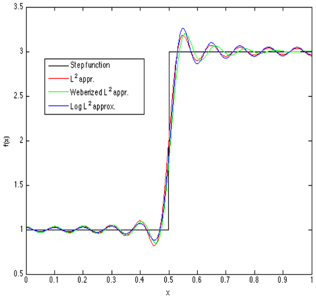
$$\begin{aligned} d_{\log}(u,v) &= d_2(\log u, \log v) \\ &= \left[\int_a^b [\log u(x) - \log v(x)]^2 \, dx\right]^{1/2} \\ &= \left[\int_a^b \left[\log \frac{v(x)}{u(x)}\right] dx\right]^{1/2} . \end{aligned} \tag{11}$$

In both of the above methods, "Weberization" is made possible by the ratio $\dfrac{v(x)}{u(x)}$.

Other efforts to incorporate Weber's model of perception in image processing methods

Example: Step function on [0,1]





Signal/image function metrics generated by greyscale range measures

These methods may be viewed as rather *ad hoc*. A more general, and mathematically-based method of defining intensity-dependent metrics between image functions is made possible considering **measures on the greyscale range space** $\mathbb{R}_g$. Such an idea was introduced in the following paper:

> *B. Forte and E.R. Vrscay, Solving the inverse problem for function and image approximation using iterated function systems, Dyn. Cont. Disc. Imp. Sys.* **1**, *177-231 (1995).*

In this paper, we wrote the following:

> *"In principle, the measure ν may be used to define various types of* greyscales, *e.g., (i) quantized grey levels, where ν consists of a finite set of Dirac measure, (ii) nonuniform distributions which model the varying sensitivities of the human eye to different regions of the grey level spectrum."*

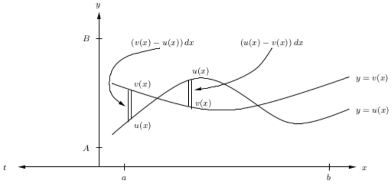
Unfortunately, we didn't know about Weber's model of perception!

Mathematical ingredients for the work which follows:

1. **Base (or pixel) space:** $X \subset \mathbb{R}$. In what follows, we let $X = [a, b]$.
2. **Greyscale range space:** $\mathbb{R}_g = [A, B] \subset (0, \infty)$. Note that this implies that all image functions are **positive-valued**.
3. **Signal/image function space:** $\mathcal{F} = \{u : X \to \mathbb{R}_g \mid u \text{ is measurable}\}$.
4. **Greyscale range measure space:** $\mathcal{M}_g$, set of probability measures on $\mathbb{R}_g$.

Goal: To assign a distance between u and v based on an integration over vertical strips of width dx centered at $x \in [a, b]$.

Generic situation:



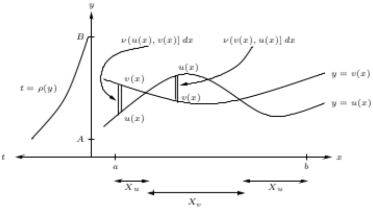
In most, if not all, traditional integration-based metrics, e.g., the $L^p(X)$ metrics for $p \geq 1$, the contribution of each strip to the integral will be an appropriate power of the height of the strip, $|u(x) - v(x)|$, i.e.,

$$d_p(u, v) = \left[\int_X |u(x) - v(x)|^p \, dx \right]^{1/p} . \tag{12}$$

This implicitly assumes a **uniform** weighting over intensity axis $\mathbb{R}_g$ since the term $|u(x) - v(x)|$ represents the Lebesgue measure of the strips.

Signal/image function metrics generated by greyscale range measures

We now wish to use measures which are **nonuniform** over $\mathbb{R}_g$. (If we keep Weber's law/model in mind, then these measures will assign lesser weight at higher intensities.)



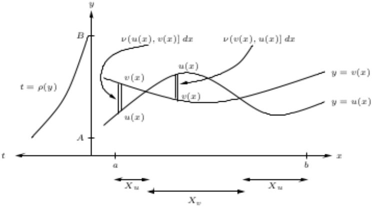
Let ν be a measure that is supported on $\mathbb{R}_g = [A, B]$. ν will now be used to define the lengths of the vertical strips as follows:

- strip at left: $\nu(u(x), v(x)]$ since $u(x) < v(x)$.
- strip at right: $\nu(v(x), u(x)]$ since $v(x) < u(x)$.

Now integrate over all strips centered at $x \in X$.

We can express this compactly as follows.

Signal/image function metrics generated by greyscale range measures



Define the following subsets of $X = [a, b]$,

$$X_u = \{x \in X \mid u(x) \le v(x)\} \qquad X_v = \{x \in X \mid v(x) \le u(x)\} \tag{13}$$

so that $X = X_u \cup X_v$. The distance between u and v associated with the measure ν is then defined as follows,

$$D(u, v; \nu) = \int_{X_u} \nu(u(x), v(x)]\, dx + \int_{X_v} \nu(v(x), u(x)]\, dx\,. \tag{14}$$

Signal/image function metrics generated by greyscale range measures

$$D(u, v; \nu) = \int_{X_u} \nu(u(x), v(x)]\, dx + \int_{X_v} \nu(v(x), u(x)]\, dx\,. \tag{15}$$

Special case: $\nu = m_g$, the usual (uniform) Lebesgue measure on $\mathbb{R}_g$, Then for $a < b$,

$$m_g(a, b] = b - a. \tag{16}$$

The sizes of the intervals shown in the above figure become

- strip at left: $\nu(u(x), v(x)] = m_g(u(x), v(x)] = v(x) - u(x)$,
- strip at right: $\nu(v(x), u(x)] = m_g(v(x), u(x)] = u(x) - v(x)$,

so that

$$\begin{aligned} D(u, v; m_g) &= \int_{X_u} [v(x) - u(x)]\, dx + \int_{X_v} [u(x) - v(x)]\, dx \\ &= \int_X |u(x) - v(x)|\, dx\,, \end{aligned} \tag{17}$$

the L^1 distance between u and v.

Signal/image function metrics generated by greyscale range measures

The natural question is, "What other kind of greyscale measures can/should be considered on $\mathbb{R}_g$?"

For convenience, we consider measures $\nu \in \mathcal{M}_g$ which are defined by continuous, non-negative density functions $\rho(y)$. (This implies that ν is absolutely continuous with respect to Lebesgue measure on $\mathbb{R}_g$.) Given a measure $\nu \in \mathcal{M}_g$ with density function ρ, then for any interval $(y_1, y_2] \subset \mathbb{R}_g$,

$$\nu(y_1, y_2) = \int_{y_1}^{y_2} \rho(y)\, dy = P(y_2) - P(y_1)\,, \tag{18}$$

where $P'(y) = \rho(y)$.

The distance function $D(u, v; \nu)$ becomes

$$\begin{aligned} D(u, v; \nu) &= \int_{X_u} \nu(u(x), v(x)]\, dx + \int_{X_v} \nu(v(x), u(x)]\, dx \\ &= \int_{X_u} [P(v(x)) - P(u(x))]\, dx + \int_{X_v} [P(u(x)) - P(v(x))]\, dx \\ &= \int_X |P(u(x)) - P(u(v))|\, dx\,. \end{aligned} \tag{19}$$

Special case: The density function

$$\rho(y) = \frac{1}{y}\,, \qquad y > 0\,. \tag{20}$$

In this case, $P(y) = \ln y$ so that distance function becomes

$$D(u, v; \nu) = \int_X |\ln u(x) - \ln v(x)|\, dx = \|\ln u - \ln v\|_1 \quad \left(= \left\| \ln\left(\frac{u}{v}\right) \right\|_1 \right)\,. \tag{21}$$

We now show that the measure associated with this density function accommodates Weber's standard model of perception.

Signal/image function metrics generated by greyscale range measures

From Kowalik-Urbaniak *et al.* (ICIAR 14):

Let $I_1, I_2 \in \mathbb{R}_g$ be any two greyscale intensities. From Weber's model, minimum changes in perceived intensity, ΔI_1 and ΔI_2, at I_1 and I_2, respectively, are given by

$$\frac{\Delta I_1}{I_1} = \frac{\Delta I_2}{I_2} = C \quad \Longrightarrow \quad \Delta I_1 = CI_1 \, , \Delta_2 = CI_2 \, . \tag{22}$$

A simple calculation shows that

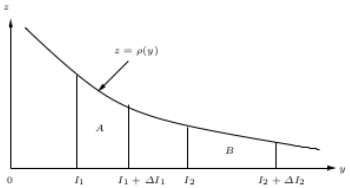
$$\int_{I_1}^{I_1+\Delta I_1} \frac{1}{y} \, dy = \int_{I_2}^{I_2+\Delta I_2} \frac{1}{y} \, dy = \ln(1+C) \Longrightarrow \nu(I_1, I_1+\Delta I_1) = \nu(I_2, I_2+\Delta I_2) \, . \tag{23}$$

This may be viewed as an invariance result with respect to perception.

Theorem: In the case $a = 1$, the density function $\rho(y) = \dfrac{1}{y}$ satisfying the above equal-area condition is unique.

Signal/image function metrics generated by greyscale range measures

Graphical interpretation in terms of equal areas enclosed by density curve $\rho(y) = \dfrac{1}{y}$:



Just to recall: The greyscale range density function $\rho(y) = \dfrac{1}{y}$ decreases with y, i.e., it assigns lesser weight to the distance integral at higher intensity values.

Generalized Weber models of perception

Given a greyscale background intensity $I > 0$, the minimum change in intensity ΔI perceived by the HVS is given by

$$\frac{\Delta I}{I^a} = C \,, \tag{24}$$

where $a > 0$ and C is constant, or at least roughly constant, over a significant range of intensities.

- The case $a = 1$ corresponds to the standard Weber model.
- There are situations in which other values of a, in particular, $a = \frac{1}{2}$ may apply

  J.A. Michon, Note on the generalized form of Weber's Law, Perception and Psychophysics **1**, *129-132 (1966).*

Density functions associated with generalized Weber models

Generalized Weber model: for a given $a > 0$ there exists a C such that

$$\frac{\Delta I}{I^a} = C \quad \Longrightarrow \quad \Delta I = CI^a\,, \tag{25}$$

where ΔI is minimum change in perceived intensity value at I.

From above, we know that for $a = 1$ there is a unique density function $\rho(y) = \dfrac{1}{y}$ sastifying the "equal-areas condition." What about the general case $a > 0$?

Definition: For a given $a > 0$, suppose that Weber's model of perception $\Delta I = CI^a$ holds for a particular value of $C > 0$ for all values of $I \geq A$. We say that a measure $\nu_a(y)$ defined by the density function $\rho_a(y)$ **conforms to** or **accomodates** this Weber model if the following (equal-area) condition holds for all $I \geq A$,

$$\nu_a(I, I + \Delta I) = \int_I^{I+\Delta I} \rho_a(y)\, dy = K\,, \tag{26}$$

for some constant $K > 0$.

In the following paper (ICIAR 18),

> *D. Li, D. La Torre and E.R. Vrscay, Image function metrics using intensity-based measures, in* Image Analysis and Recognition, *ICIAR 18, LNCS* **10882**, *326-335 (2018).*

we showed that for $0 < a < 1$, if a density $\rho_a(y)$ accomodating Weber's model, then

$$\rho_a(y) \simeq \frac{1}{y^a} \quad \text{as} \quad y \to \infty \, . \tag{27}$$

The proof employed an ODE that is associated with the equal-area condition.
(In the case $a = 1$, the equation is exact.)

However, we did not prove existence/uniqueness. In what follows, we prove existence and uniqueness.

The following result, although quite trivial, will have important consequences.

Theorem: For given values of $a > 0$ and $C > 0$, let $\nu_a(y)$ be a measure with density function $\rho_a(y)$ which conforms to Weber's model according to Eq. (26) above. Then $\nu(A, \infty) = \displaystyle\int_A^\infty \rho_a(y)\, dy = \infty$.

Sketch of Proof: Let $y_0 = A$ and $y_{n+1} = y_n + Cy_n^a$ for $n \geq 0$. It is not difficult to show that $y_n \to \infty$ as $n \to \infty$. From Eq. (26),

$$\int_A^{y_n} \rho_a(y)\, dy = nK \to \infty \quad \text{as} \quad n \to \infty\,. \tag{28}$$

The main result:

Theorem: For given values of $a > 0$ and $C > 0$, there exists a unique, continuous function $\rho_a(y)$ defined on $[A, \infty)$ which satisfies the equal-area Weber condition in Eq. (26).

Main ideas behind the proof of this Theorem:

Let $g(y)$ be a continuous function on $[A, \infty)$ and define

$$G(x) = \int_A^x g(y)\, dy \quad x \geq A\,. \tag{29}$$

Clearly $G(A) = 0$. Furthermore, suppose that for fixed values of $a > 0$ and $C > 0$, $g(y)$ satisfies Eq. (26). This means that

$$G(x + Cx^a) - G(x) = K\,, \quad x \geq A\,. \tag{30}$$

For convenience, define $f(x) = x + Cx^a$ and divide both sides of the above equation by K to obtain

$$H(f(x)) - H(x) = 1\,, \tag{31}$$

where $H(x) = K^{-1}G(x)$. Eq. (31) is known as **Abel's equation**, a well-known **functional equation** which is important in the theory of iteration.

Now consider the following linear functional equation for $g(x)$ for $x \geq A$,

$$g(x + Cx^a)(1 + aCx^{a-1}) - g(x) = 0 \,. \tag{32}$$

(This equation may be obtained by differentiating both sides of Eq. (26), where $\rho(y) = g(y)$, with the assumption that $g(y)$ is continuous.) Eq. (32) is a special case of the following family of linear functional equations studied by Belitskii and Lyubich*,

$$P(x)\psi(x)(F(x)) + Q(x)\psi(x) = \gamma(x) \quad x \in X \,, \tag{33}$$

where X is the topological space over which the equation is being considered. Here, $X = [A, \infty)$, $P(x) = 1 + aCx^{a-1}$, $F(x) = 1 + aCx^a$, $Q(x) = -1$ and $\gamma(x) = 0$. In [1] it is shown that if the Abel equation associated with Eq. (32), namely,

$$\phi(F(x)) - \phi(x) = 1 \,, \tag{34}$$

has a continuous solution $\phi(x)$, then Eq. (32) is **totally solvable**, i.e., it has a continuous solution $\psi(x)$ for every continuous function $\gamma(x)$.

*G. Belitskii and Y. Lyubich, The Abel equation and total solvability of linear functional equations, Studia Mathematica **127** (1), 81-97 (1998).

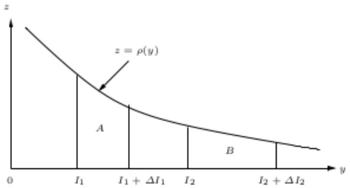
The existence of continuous solutions to Eq. (34) depends on the iteration dynamics of the function $F(x)$ on the space X. In this case, $F(x) = x + Cx^a$ is an increasing function on $X = [A, \infty)$ and for all $x \in X$, $F^n(x) \to \infty$ as $n \to \infty$. As such, the conditions of Corollary 1.6 in the BL paper are satisfied, i.e., every compact set $S \subset X$ is wandering under the action of F. Therefore, Eq. (34) has a continuous solution (unique up to a constant). This, in turn, implies that Eq. (32) has a unique, nonzero, continuous solution.

Additional analysis of Eqs. (26) and (32) yields the following properties of $\rho_a(y)$ for $a > 0$ which we state without formal proof:

- $\rho_a(y) > 0$ for all $y > 0$. (Or at least we choose a positive solution of Eq. (32). If $g(x)$ is a solution to Eq. (32), then so is $Cg(x)$ for any $C \in \mathbb{R}$.
- $\rho_a(y)$ is decreasing on $[0, \infty)$ and $\rho_a(y) \to \infty$ as $y \to \infty$. This is to be expected: In Eq. (26), the length of the interval $y + cy^a$ increases with y. The equal-area condition dictates that ρ_a decrease with y.
- As $y \to 0^+$, $\rho_a(y) \to \infty$. This is also expected from the equal-area condition since the length of the interval $[x, x + Cx^a]$ decreases as $x \to 0^+$.

Density functions associated with generalized Weber models

Graphical interpretation in terms of equal areas enclosed by (unique) density curve $\rho_a(y)$ which accommodates Weber's model for $a > 0$



Once again: For $a > 0$, the greyscale range density function $\rho_a(y)$ decreases with y, i.e., it assigns lesser weight to the distance integral at higher intensity values.

The above invariance/equal-area result may also be extended to include the special case $a = 0$, i.e.,

$$\frac{\Delta I}{I^0} = C \quad \Longrightarrow \quad \Delta I = C\,, \tag{35}$$

essentially an absence of Weber's model. In this case,

$$\rho_0(y) = 1\,, \tag{36}$$

which corresponds to **uniform Lebesgue measure** m_g.

Asymptotic behaviour of density functions $\rho_a(y)$ which conform to Weber's model

The determination of the asymptotic behaviour of the density functions $\rho_a(y)$ is centered on the equal-area property of Eq. (26). Here we simply state the results which are obtained, for the most part, using standard Calculus and algebraic manipulation.

Asymptotic behaviour as $y \to \infty$

- $0 < a < 1$:

$$\rho_a(y) = \frac{1}{y} + \sum_{n=1}^{\infty} \frac{A_n}{y^{n(1-a)}} \tag{37}$$

- $a = 1$:

$$\rho_a(y) = \frac{1}{y} \tag{38}$$

- $a > 1$:

$$\rho_a(y) \simeq \frac{1}{y \ln y} - \left(\frac{\ln C}{a}\right) \frac{1}{y(\ln y)^2} \tag{39}$$

Asymptotic behaviour of density functions $\rho_a(y)$ which conform to Weber's model

There is a reciprocity with regard to integrals involved in the above analysis of $y \to \infty$ and those involved in the case $y \to 0^+$. For example, an analysis of the limit $y \to 0^+$ for the case $0 < a < 1$ employs the same equations as those used in the analysi of the limit $y \to \infty$ in the case $a > 1$. Net result:

Asymptotic behaviour as $y \to 0^+$

- $0 < a < 1$:

$$\rho_a(y) \simeq \frac{1}{y \ln y} - \left(\frac{\ln C}{a}\right) \frac{1}{y(\ln y)^2} \tag{40}$$

- $a = 1$:

$$\rho_a(y) = \frac{1}{y} \tag{41}$$

- $a > 1$:

$$\rho_a(y) = \frac{1}{y} + \sum_{n=1}^{\infty} \frac{A_n}{y^{n(1-a)}} \tag{42}$$

Distance functions associated with generalized Weber model density functions

For simplicity, we consider only the leading order behaviour of the density functions $\rho_a(y)$.

For $a < 0 < 1$, we then simply define

$$\rho_a(y) = \frac{1}{y^a} \quad \Longrightarrow \quad P(y) = \frac{1}{-a+1} y^{-a+1} \,. \tag{43}$$

Associated distance functions, up to a multiplicative constant, are given by

$$D_a(u,v) = D(u,v;\nu_a) = \int_X \left| u(x)^{-a+1} - v(x)^{-a+1} \right| \, dx \,. \tag{44}$$

For $a = 1$ (Weber's standard model),

$$\rho_1(y) = \frac{1}{y} \quad \Longrightarrow \quad P(y) = \ln y \,. \tag{45}$$

Associated distance function, up to a multiplicative constant,

$$D_1(u,v) = D(u,v;\nu_a) = \int_X |\ln u(x) - \ln v(x)| \, dx \,. \tag{46}$$

Function approximation using generalized greyscale measure ν_a

Let $u(x)$ denote a reference function and $v(x)$ an approximation to $u(x)$ having standard form,

$$v(x) = \sum_{k=1}^{N} c_k \phi_k(x) \,, \tag{47}$$

where the set $\{\phi_k\}_{k=1}^{N}$ is assumed to be linearly independent, and perhaps orthogonal, over $X = [a, b]$. Let $Y_N = \text{span}\{\phi_1, \phi_2, \cdots, \phi_N\}$.

Best Y_N-approximation of $u \in \mathcal{F}$ in the metric space $(\mathcal{F}, D_a)$ is found by minimizing the distance $D(u, v; \nu_a)$.

Unfortunately, it is difficult to work with these distance functions, especially because of the appearance of the absolute value in integrand.

It is easier to work with their L^2 analogues (which are also metrics):

Function approximation using L^2 analogues of distance functions associated with generalized Weber measures ν_a

For the case $a \geq 0$, $a \neq 1$:

$$D_{2,a}(u, v; \nu_a) = \left[\int_X \left[u(x)^{-a+1} - v(x)^{-a+1} \right]^2 \, dx \right]^{1/2} . \tag{48}$$

And for the case $a = 1$:

$$D_{2,1}(u, v; \nu_a) = \left[\int_X [\ln u(x) - \ln v(x)]^2 \, dx \right]^{1/2} . \tag{49}$$

Special case: $a = \dfrac{1}{2}$

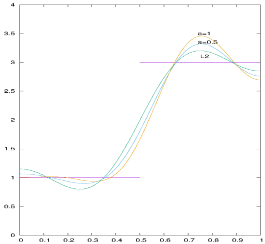
$$D_{2,1/2}(u, v; \nu_a) = \left[\int_X \left[\sqrt{u(x)} - \sqrt{v(x)} \right]^2 \, dx \right]^{1/2} . \tag{50}$$

(See, I told you!)

In all cases, gradients can be computed $\Longrightarrow$ gradient descent methods can be used.

Function approximation using L^2 analogues of distance functions associated with generalized Weber measures ν_a

Example: Return to step function on [0,1]



Best approximations to step function $u(x)$ using $N = 5$ orthogonal cosine functions on [0,1] using three greyscale measures ν_a: (i) $a = 0$ (best L^2, green), (ii) $a = 0.5$ (blue) and (iii) $a = 1$ (standard Weber, yellow).

Note that as a increases, the density function $y_a = \dfrac{1}{y^a}$ assigns less and less weight to higher intensity regions. As a result, the approximations are poorer as a increases at higher intensity regions, better as a increases at lower intensity regions.

Function approximation using L^2 analogues of distance functions associated with generalized Weber measures ν_a

Example: 2D step function on $[0,1]^2$

256×256-pixel 8 bpp image composed of four squares with greyscale values 60, 128, 128 and 220.

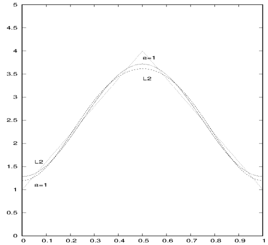
 

Left and right, respectively: Best approximations for $a = 0$ (best L^2) and $a = 1$ (standard Weber) obtained with 2D DCT basis set $\Phi_{kl}(n,m) = \phi_k(n)\phi_l(m)$, $0 \le k, l \le 14$. The $a = 1$ approximation exhibits greater deviation at higher greyscale levels than the L^2 approximation.

Function approximation using L^2 analogues of distance functions associated with generalized Weber measures ν_a

Example: 1D ramp function on $[0, 1]$

256×256-pixel 8 bpp image



Function approximation using L^2 analogues of distance functions associated with generalized Weber measures ν_a

Example: 2D ramp function on $[0,1]^2$

256×256-pixel 8 bpp image

Left: Best L^2 approximation ($a = 0$). **Center:** 2D ramp function. **Right:** Best Weberized L^2 approximation ($a = 1$).