B. D. SHARMA
Y. D. MATHUR
J. MITTER

 Bounds on the rate-distortion function for geometric measure of distortion


<http://www.numdam.org/item?id=M2AN_1973__7_2_29_0>

© AFCET, 1973, tous droits réservés.
BOUNDS ON THE RATE-DISTORTION FUNCTION FOR GEOMETRIC MEASURE OF DISTORTION

by B. D. SHARMA, Y. D. MATHUR and J. MITTER (*)

Abstract. — Earlier the authors have defined the Geometric Measure of Distortion \( aD_g \) where \( a(>0) \) stands for the cost for distortion per letter for correct transmission. In this paper we calculate the Rate Distortion Function \( R(aD_g) \). In Section 3, the Symmetric Measure of Distortion is defined and bounds are obtained on \( R(aD_g) \) and \( aD_g \).

1. INTRODUCTION

In a communication process, let \( \{ x_i \}_{i=0}^{N-1} \) be the set of symbols transmitted and \( \{ Y_j \}_{j=0}^{M-1} \) be the set of symbols received such that for correct transmission \( x_i \) corresponds to \( y_j \) for every \( i \). For an independent letter source, we shall denote by \( p_i \), the probability of transmitting \( x_i \); and by \( q_{ji} \), the probability of receiving \( y_j \) when \( x_i \) is sent. The average mutual information is given by

\[
I(P;Q) = \sum_{i=0}^{N-1} \sum_{j=0}^{M-1} p_i q_{ji} \log \left( \frac{q_{ji}}{\sum_{i} p_i q_{ji}} \right)
\]  

(1.1)

For convenience, the logarithms are considered to the base \( e \). For a transmission with a fidelity criterion [3], the authors [4] have introduced the geometric measure of distortion given by

\[
aD_g = \prod_{i,j} \rho_{ij}^{p_i q_{ji}},
\]  

(1.2)

where \( \rho_{ij} \) is the distortion (cost) of transmitting \( x_j \) and receiving \( y_j \) so that

\[
\rho_{ij} > \alpha \quad \text{if} \quad i \neq j \quad \text{and} \quad \rho_{ii} = \alpha \quad \text{where} \quad \alpha > 0
\]  

(1.3)

(1) Faculty of Mathematics University of Delhi (India).

The rate distortion function of the source relative to the given distortion measure is then defined as

\[ R(\alpha D^* G) = \min I(P; Q), \]  

(1.4)

where the minimization is done with respect to \( q_{ji} \) under the condition that

\[ \alpha D_G \leq \alpha D^*_G. \]

(1.5)

Gallager [2]; Berger [1] and others have investigated noisy channel coding theorems with the Shannon's measure of distortion given by

\[ D_S = \sum_i \sum_j p_i \cdot q_{ji} \cdot d_{ij}, \]

(1.6)

in which \( d_{ij} > 0 \) if \( i \neq j \) and \( d_{ii} = 0 \) (1.7)

In this paper, we shall investigate the values of \( R(\alpha D^*_G) \) and prove theorems on the symmetric measure of distortion with the geometric fidelity criterion.

It is rather obvious that \( R(\alpha D^*_G) \) is non-negative and a non-increasing function of \( \alpha D_G \) for minimization in (1.4) is done over a constraint set which is enlarged as \( \alpha D^*_G \) is increased.

2. CALCULATION OF \( R(\alpha D^*_G) \)

**Theorem 2.1** The set \( \{ q_{ji} \} \) which gives \( R(\alpha D^*_G) \) i.e. \( \min I(P; Q) \) subject to the constraint \( \alpha D_G \leq \alpha D^*_G \) is given by

\[ q_{ji} = \frac{q_j \cdot c_i \cdot \rho_{ij}^{-\lambda \alpha D_G}}{p_i} \quad \text{for all } i, j \]

(2.1)

where

\[ \sum_i c_i \rho_{ij}^{-\lambda \alpha D_G} = 1 \quad \text{for all } j \quad \text{and} \quad q_j = \sum_i p_i \cdot q_{ji}. \]

(2.2)

**Proof:** We have to minimize (1.1) under the conditions

\[ \alpha D_G = \exp \left( \sum_i \sum_j p_i \cdot q_{ji} \cdot \log \rho_{ij} \right) \leq \alpha D^*_G \]

and \( \sum_j q_{ji} = 1 \) for all \( i \).

Consider the function

\[ \Phi = I(P; Q) + \lambda \cdot \alpha D_G + \sum_i \mu_i \cdot \sum_j q_{ji} \]

(2.3)

where \( \lambda \) and \( \mu_i \) are Lagrange's constants.

For a suitable choice let \( \mu_i = -p_i \log \frac{c_i}{p_i} \)

(2.4)
Replacing the set $\mu = \{ \mu_i \}_{i=0}^{N-1}$ by $c = \{ c_i \}_{i=0}^{N-1}$, (2.3) becomes

$$
\Phi = \sum_{i=0}^{N-1} \sum_{j=0}^{M-1} p_i \cdot q_{ji} \left( \log \frac{q_{ji}}{\sum_j p_j \cdot q_{ji}} - \log \frac{c_i}{p_i} \right) + \lambda \cdot \exp \left( \sum_i \sum_j p_{ij} \cdot \log \rho_{ij} \right)
$$

(2.5)

Thus the condition for $q_{ji}$ to yield a stationary point for $\Phi$ is

$$
\log \frac{q_{ji}}{q_j} + \lambda \cdot \exp \left( \sum_i \sum_j p_i q_{ji} \log \rho_{ij} \right) \log \rho_{ij} - \log \frac{c_i}{p_i} = 0
$$

(2.6)

for every $i$ and $j$

where

$$
q_j = \sum_i p_i \cdot q_{ji}
$$

(2.7)

Next (2.6) gives

$$
q_{ji} = \frac{c_i}{p_i} \cdot q_j \cdot \rho_{ij}^{-\lambda \cdot aD\alpha}.
$$

(2.8)

Multiplying (2.8) by $p_i$ and summing over $i$, we get

$$
\sum_i c_i \cdot \rho_{ij}^{-\lambda \cdot aD\alpha} = 1 \quad \text{for all } j.
$$

(2.9)

Again summing up (2.8) over $j$ and using the constraint $\sum_j q_{ji} = 1$ for every $i$, we obtain

$$
\frac{c_i}{p_i} \sum_j q_j \cdot \rho_{ij}^{-\lambda \cdot aD\alpha} = 1 \quad \text{for all } i.
$$

(2.10)

From (2.9) we get a set of $M$-linear equations in the unknowns $c_i$ and another set of $M$-linear equations in $q_j$ obtained from (2.10). If $N = M$, we can usually solve the equations and then find $q_{ji}$ from (2.8). Since $I(P : Q)$ is convex in $Q$, $\Phi$ is also convex and therefore the solution is a minimum.

The above approach does not take into account the non-negativity of quantities $q_{ji}$ and the resulting values of $q_{ji}$, giving minimum of $I(P : Q)$ may become negative, leading to a non-feasible solution.

In the next theorem we follow an approach which always gives a feasible solution.

Now we define a function

$$
\psi = \sum_i \sum_j p_i \cdot q_{ji} \left[ \log \frac{q_{ji}}{\sum_j p_j \cdot q_{ji}} \right] + \lambda \cdot D_G
$$

(2.11)

where $q_{jj} > 0$. 

It would be noted that since $\alpha D_G \leq \alpha D_G^*$

$$\min \psi \rightarrow \lambda \cdot \alpha D_G^* \leq R(\alpha D_G^*). \quad (2.12)$$

**Theorem 2.2** For any $\lambda > 0$,

$$\min \psi = H(U) + \max_{q_{ij}} \sum_{i=0}^{N-1} p_i \cdot \log c_i; \quad c_i > 0, \quad (2.13)$$

where $H(U)$ is the entropy of the source and $C = \{ c_i \}_{i=0}^{N-1}$ is such that

$$\sum_{i=0}^{N-1} c_i \cdot \rho_{ij}^{-\lambda \cdot aD_G} \leq 1 \quad \text{where} \quad \rho_{ij} \geq e. \quad (2.14)$$

Also $\psi$ is minimized for values of $c_i$ given by (2.8) in terms of $q_{ij}$ and the necessary and sufficient conditions on $c_i$ to achieve the maximum in (2.13) are that there exists an output distribution satisfying (2.10) and (2.14) with equality.

**Proof:** Consider the function

$$\Phi = \sum_i \sum_j p_i \cdot q_{ji} \cdot \log \frac{q_{ji}}{q_j} + \lambda \cdot \alpha D_G - \sum_i p_i \log \frac{c_i}{p_i} \sum_j q_{ji} \quad (2.15)$$

then

$$\Phi = \psi - H(U) - \sum_i p_i \cdot \log c_i \quad (2.16)$$

(2.15) can be put as

$$-\Phi = \sum_i \sum_j p_i q_{ji} \log \frac{q_{ji} \cdot c_i}{q_{ji} \cdot p_i} + \sum_i \sum_j p_i \cdot q_{ji} \cdot \log \rho_{ij}^{-\lambda \cdot aD_G} \cdot \log \rho_{ij} e$$

$$\leq \sum_i \sum_j p_i \cdot q_{ji} \log \frac{q_{ji} \cdot c_i}{q_{ji} \cdot p_i} + \sum_i \sum_j p_i \cdot q_{ji} \cdot \log \rho_{ij}^{-\lambda \cdot aD_G} \quad \text{as} \quad \rho_{ij} \geq e. \quad (2.17)$$

Using the inequality $\log x \leq x - 1$, we obtain

$$-\Phi \leq \sum_i \sum_j p_i q_{ji} \left[ \frac{q_{ji} \cdot c_i \cdot \rho_{ij}^{-\lambda \cdot aD_G}}{q_{ji} \cdot p_i} - 1 \right]$$

$$= \sum_i \sum_j q_j \cdot c_i \cdot \rho_{ij}^{-\lambda \cdot aD_G} - \sum_j q_j$$

$$\leq \sum_j q_j - \sum_j q_j = 0 \quad (2.17)$$

(Using (2.14))

Combining (2.17) and (2.16), we get

$$\psi \geq H(U) + \sum_i p_i \log c_i \quad (2.18)$$
(2.18) is satisfied with equality if and only if the inequalities \( \log x < x - 1 \) and (2.14) are satisfied with equality, or if and only if

\[
\frac{q_j \cdot c_i \cdot p_{ij}^{-\lambda D \alpha}}{q_{ji} \cdot p_i} = 1 \quad \text{for all} \quad q_{ji} > 0
\]

(2.19)

and

\[
\sum_i c_i p_{ij}^{-\lambda D \alpha} = 1 \quad \text{for all} \quad q_j > 0
\]

(2.20)

The conditions in the theorem are necessary for equality in (2.18) as we obtain (2.10) from (2.19) after multiplying by \( q_{ji} \) and summing over \( j \). Again if the output probabilities satisfy (2.10) and if (2.20) is satisfied then as already seen \( q_{ji} \) given by (2.8) is a transition assignment with output probabilities \( q_j \). By (2.10), the choice satisfies (2.19) so that the conditions of the theorem are sufficient for equality in (2.18).

3. SYMMETRIC MEASURE OF DISTORTION

If the number of input and output symbols are same and if the cost of correct transmission is \( \alpha \) and the cost of any incorrect transmission is \( \beta \) (obviously \( \alpha < \beta \)) so that the distortion is

\[
P_{ij} = \begin{cases} 
\alpha & \text{if } i = j \\
\beta & \text{if } i \neq j 
\end{cases}
\]

(3.1)

then we refer to this as Symmetric Measure of Distortion.

**Theorem 3.1.** Under symmetric measure of Distortion, we have

\[
R_{(a D^* G)} \geq H(U) - \frac{1}{\beta} \left( \frac{\alpha D^* G - \alpha \log \alpha}{\beta \log \beta - \alpha \log \alpha} \right) \log(N - 1)
\]

(3.2)

where

\[
H \left( \frac{\alpha D^* G - \alpha \log \alpha}{\beta \log \beta - \alpha \log \alpha} \right) = \left( \frac{\alpha D^* G - \alpha \log \alpha}{\beta \log \beta - \alpha \log \alpha} \right) \log \left( \frac{\alpha D^* G - \alpha \log \alpha}{\beta \log \beta - \alpha \log \alpha} \right) - \left( \frac{\alpha D^* G - \alpha \log \alpha}{\beta \log \beta - \alpha \log \alpha} \right) \log \left( 1 - \frac{\alpha D^* G - \alpha \log \alpha}{\beta \log \beta - \alpha \log \alpha} \right)
\]

(3.3)

with equality if

\[
\alpha D^* G \leq \alpha \log \alpha + \left( \beta \log \beta - \alpha \log \alpha \right) (N - 1) p_{\min},
\]

where \( p_{\min} \) is the minimum of all \( p_i \)'s.

\( \text{n° août 1973, R-2.} \)
Proof: The constraint equations (2.14) under symmetric measure of distortion take the form

\begin{equation}
C_j \cdot \alpha^{-\lambda} + \left( \sum_{i=0}^{N-1} c_i - c_j \right) \beta^{-\lambda} \leq 1 \quad (3.5)
\end{equation}

\[ 0 \leq j \leq M - 1. \]

These are all symmetric and can be made to hold with equality by taking \( c_i = c_0 \) for each \( i \). Then,

\[ c_0 = \alpha^{\lambda} \cdot [1 + (N - 1) \cdot \alpha^{\lambda} \cdot \beta^{-\lambda}]^{-1} \quad (3.6) \]

From (2.13) and (3.6), we have

\[ \min \psi \geq H(U) + \lambda \cdot \alpha \log \alpha - \log [1 + (N - 1) \alpha^{\lambda} \cdot \beta^{-\lambda}] \quad (3.7) \]

Invoking the relation (2.12) we get for all \( \lambda > 0 \),

\[ R(\alpha D_G^*) \geq -\lambda \cdot \alpha D_G^* + H(U) + \lambda \cdot \alpha \log \alpha - \log [1 + (N - 1) \alpha^{\lambda} \cdot \beta^{-\lambda}] \quad (3.8) \]

Now if we maximize the right hand side with respect to \( \lambda \), we get

\[ \alpha D_G^* = \alpha \log \alpha + \frac{(\beta \log \beta - \alpha \log \alpha)(N - 1)}{\beta^{\lambda} \cdot \alpha^{-\lambda} + (N - 1)} \quad (3.9) \]

therefore

\[ \lambda = \frac{1}{\beta \log \beta - \alpha \log \alpha} \log \left( \frac{\beta \log \beta - \alpha D_G^*}{\alpha D_G^* - \alpha \log \alpha} \right)(N - 1). \quad (3.10) \]

(3.2) follows by substituting (3.10) into (3.8).

Now by theorem 2.2 (3.7) would hold with equality if we can find a solution of (2.10) such that \( q_i \geq 0 \). Under the symmetric measure of distortion defined by (3.1), (2.10) gives

\[ q_j = \frac{(p_{i,\alpha}) \cdot \alpha^{\lambda} \cdot \beta^{\lambda} - \alpha^{\lambda}}{\beta^{\lambda} - \alpha^{\lambda}} \quad (3.11) \]

\[ = p_i [\beta^{\lambda} + (N - 1) \alpha^{\lambda}] - \alpha^{\lambda} \quad (3.12) \]

for values of \( c_i = c_0 \) given in (3.6).

All \( q_j \)'s will be non negative if

\[ p_i \geq \frac{1}{\beta^{\lambda} \cdot \alpha^{-\lambda} + (N - 1)} \quad \text{for every } i \quad (3.13) \]
If \( \lambda \) is sufficiently large (3.13) holds normally and (3.7) would hold with equality.

Now combining (3.9) and (3.13), we get

\[
R(\lambda^D_\beta) = H(U) - H\left(\frac{\alpha D^*_\beta - \alpha \log \alpha}{\beta \log \beta - \alpha \log \alpha}\right) - \left(\frac{\alpha D^*_\beta - \alpha \log \alpha}{\beta \log \beta - \alpha \log \alpha}\right) \log (N - 1)
\]

for

\[
\alpha D^*_\beta \leq \alpha \log \alpha + (\beta \log \beta - \alpha \log \alpha)(N - 1)p_{\min}.
\]

Hence the theorem.

**An Extension of Theorem 3.1**

We shall now calculate \( R(\lambda^D_\beta) \) for large values of \( \lambda^D_\beta \). Without any loss of generality we can assume that the source letters are ordered in decreasing order of probabilities that is

\[
p_0 \geq p_1 \geq \ldots \geq p_{N-1}.
\]  

(3.14)

Next suppose that there is an integer \( m, 0 < m < N - 1 \) such that

\[
q_j = \begin{cases} 
0 & \text{if } j \geq m \\
> 0 & \text{if } j \leq m - 1.
\end{cases}
\]

(3.15)

For \( j \leq m \), (3.11) then gives

\[
p_i = c_i \beta^{-\lambda \beta}.
\]

(3.16)

(3.5) must be satisfied with equality for \( j \leq m - 1 \), therefore for all \( j \leq m - 1 \), all the \( c_j \) must be the same say \( c_0 \) and \( c_j \leq c_0 \) for \( j > m \).

The constraint equations (2.14) for \( j = 0 \) gives

\[
c_0 \alpha^{-\lambda \alpha} + \left(\sum_{i=0}^{m-1} c_i + \sum_{i=m}^{N-1} c_i - c_0\right) \cdot \beta^{-\lambda \beta} = 1,
\]

or

\[
c_0 \alpha^{-\lambda \alpha} + (mc_0 - c_0) \beta^{-\lambda \beta} + \sum_{i=m}^{N-1} c_i \cdot \beta^{-\lambda \beta} = 1, \quad (3.17)
\]

or

\[
c_0 \left[\alpha^{-\lambda \alpha} + (m - 1) \beta^{-\lambda \beta}\right] = \sum_{i=0}^{m-1} p_i = \sigma_m(\text{say})
\]

(using (3.16))

or

\[
c_i = c_0 = \frac{\sigma_m \cdot \alpha^{\lambda \alpha}}{1 + (m - 1) \alpha^{\lambda \alpha} \cdot \beta^{-\lambda \beta}}
\]

(3.18)

It is clear from (3.16) that \( c_m \geq c_{m+1} \geq \ldots \geq c_{N-1} \) and for \( j \geq m \); \( c_j \leq c_0 \) will hold if

\[
P_m \leq \frac{\sigma_m \cdot \alpha^\lambda \cdot \beta^{-\lambda \beta}}{1 + (m - 1) \alpha^\lambda \cdot \beta^{-\lambda \beta}}
\]

(3.19)

Now \( \sum_i p_i \log c_i \) will be maximized for \( c \) given by (3.16) and (3.18), if all the \( q_j \)'s given by (3.11) are non negative. This requires from (3.11) that

\[
P_{m-1} \geq \frac{\sigma_m \cdot \alpha^\lambda \cdot \beta^{-\lambda \beta}}{1 + (m - 1) \alpha^\lambda \cdot \beta^{-\lambda \beta}}
\]

(3.20)

since from (3.11) and (3.14) it is obvious that

\[
q_0 \geq q_1 \geq \ldots \geq q_{m-1}.
\]

Thus for the values of \( \lambda \) for which (3.19) and (3.20) are satisfied, the given \( c \) yields

\[
\min \psi = H(U) + \sum_{i=0}^{m-1} p_i \log \frac{\sigma_m \cdot \alpha^\lambda}{1 + (m - 1) \alpha^\lambda \cdot \beta^{-\lambda \beta}} + \sum_{i=m}^{N-1} p_i \log (p_i \cdot \beta^\lambda).
\]

(3.21)

The min \( \psi \) over a range of \( \lambda \) specifies \( R(\alpha D^*_G) \) over the corresponding range of \( \lambda \). The parameter \( \lambda \) is related to \( \alpha D^*_G \) by

\[
\alpha D^*_G = \frac{\partial}{\partial \lambda} \left[ \min \psi \right] = \sigma_m \left[ \frac{\alpha \log \alpha + (\beta \log \beta)(m - 1) \alpha^\lambda \beta^{-\lambda \beta}}{1 + (m - 1) \alpha^\lambda \beta^{-\lambda \beta}} \right] + (\beta \log \beta)(1 - \sigma_m).
\]

(3.22)

Therefore

\[
\lambda = \log \left[ \frac{(m - 1)(\beta \log \beta - \alpha D^*_G)}{\alpha D^*_G - \beta \log \beta + (\beta \log \beta - \alpha \log \alpha)\sigma_m} \right]^{1/(\beta \log \beta - \alpha \log \alpha)}
\]

(3.23)

For \( \lambda \) and \( \alpha D^*_G \) related by (3.22).

\[
R(\alpha D^*_G) = \min \psi - \lambda \cdot \alpha D^*_G
\]

(3.24)

using (3.21) and (3.23); simplifying and rearranging the terms, (3.24) becomes

\[
R(\alpha D^*_G) = \sigma_m \left[ H(U_m) + \left\{ \frac{\alpha D^*_G - \beta \log \beta + (\beta \log \beta - \alpha \log \alpha)\sigma_m}{(\beta \log \beta - \alpha \log \alpha)\sigma_m} \right\} \times \log \left\{ \frac{\alpha D^*_G - \beta \log \beta + (\beta \log \beta - \alpha \log \alpha)\sigma_m}{(\beta \log \beta - \alpha \log \alpha)\sigma_m} \right\} \right]
\]

Revue Française d'Automatique, Informatique et Recherche Opérationnelle
BOUNDS ON THE RATE-DISTORSION FUNCTION

\[ R(\alpha D^*_G) = \sigma_m[H(U_m) - \hat{H}(\Delta) - \Delta \log (m - 1)] \]

This can be equivalently expressed as

\[ R(\alpha D^*_G) = \sigma_m[H(U_m) - \hat{H}(\Delta) - \Delta \log (m - 1)] \]

where \( H(U_m) \) is the entropy of a reduced ensemble with probabilities

\[ p_0/\sigma_m, \quad p_1/\sigma_m, \ldots, p_{m-1}/\sigma_m. \]

\[ \Delta = \frac{\beta \log \beta - \alpha \log \alpha}{(\beta \log \beta - \alpha \log \alpha) \sigma_m} \]

and

\[ \hat{H}(\Delta) = -\Delta \log \Delta - (1 - \Delta) \log (1 - \Delta). \]

Substituting (3.23) into (3.19) and (3.20) we obtain the bounds of \( \alpha D^*_G \) given by

\[ (\beta \log \beta - \alpha \log \alpha) \left( mp_m - \sum_{i=0}^{m} p_i \right) + \beta \log \beta \leq \alpha D^*_G \leq (\beta \log \beta - \alpha \log \alpha) \times \left[ (m - 1)p_{m-1} - \sum_{i=0}^{m-1} p_i \right] + \beta \log \beta. \]

When \( m = N - 1 \)

\[ (\beta \log \beta - \alpha \log \alpha) \left( mp_m - \sum_{i=0}^{m} p_i \right) = \alpha \log \alpha + (\beta \log \beta - \alpha \log \alpha)(N - 1)p_{\min}. \]

which is the same as upper limit in (3.4).

ACKNOWLEDGEMENTS

The authors are thankful to Prof. U. N. Singh, Dean, Faculty of Mathematics, University of Delhi, and the third author is thankful to C.S.I.R. (India) for financial assistance.
Shannon introduced $\rho_{ij}$ as the single letter distortion when $x_i$ is sent and
$y_j$ is received. As there is always some cost even for correct transmission, we
take $\rho_{ij} > \alpha$ for $i \neq j; \alpha > 0$ and $\rho_{ij} = \alpha$ (where $\alpha$ is zero in Shannon's case).
Since any measure of distortion is an average of per letter distortions $\rho_{ij}$'s,
the measure in its most-generalized form is taken as

$$\alpha D_{\psi} = \psi^{-1} \left( \frac{\sum_i \sum_j f(p_{ij}) \psi(\rho_{ij})}{\sum_i \sum_j f(p_{ij})} \right)$$

where (i) $\psi$ is strictly monotonic and continuous function defined for non
negative values.

and (ii) $f$ is positive valued and bounded weight function in $[0, 1]$

By setting $f(x) = x$ and $\psi(x) = \log x$ in (A) we get

$$\alpha D_G = \exp \left( \sum_i \sum_j p_{ij} \cdot \log \rho_{ij} \right) = \prod_{i,j} \rho_{ij}^{p_{ij}/\alpha} \quad \text{where} \quad \sum_i \sum_j p_{ij} = 1$$

(*) For relevant matter of [4] refer to Appendix.