Vanishing Distortion and Shrinking Cells

Andrew B. Nobel *

December 18, 1995

Abstract

We establish an asymptotic connection between vanishing r'th power distortion and shrinking cell diameters for vector quantizers with convex cells.

Appears in IEEE Transaction on Information Theory,
vol. 24, no. 3, pp. 1303-1305, 1996.

*Andrew Nobel is with the Department of Statistics, University of North Carolina, Chapel Hill, NC 27599-3260. This work was completed while he was a Beckman Institute Fellow at the Beckman Institute for Advanced Science and Technology, University of Illinois, Urbana-Champaign. This work was supported in part by NSF grant DMS-9501926.

Key words and phrases. Asymptotic quantization theory, high-rate quantization theory, vector quantization, shrinking cell diameters, convex sets.

1 Introduction

The study of high rate vector quantization is concerned with the performance of quantizers having large codebooks. One seeks bounds (or precise estimates) for the asymptotic distortion of a sequence vector quantizers in terms of codebook size, vector dimension, and the nature of the selected distortion measure. Gersho [4], Yamada *et al.* [14], and others have given heuristic derivations of formulae governing the distortion of vector quantizers with large codebooks. They assume that the underlying distribution has a smooth density, and that the cell diameters of the n'th quantizer tend to zero as n tends to infinity. The latter condition, stipulating shrinking cells, is the subject of this paper.

Recently, several authors [11, 5, 13, 12, 6, 10] have proposed using vector quantizers as the basis for multivariate histogram classification and regression schemes in higher dimensions. Verification of shrinking cell conditions is typically the key to establishing the consistency of such schemes (cf. [6, 10]). Although they do not figure explicitly in the rigorous derivation [15, 16, 2, 3] of bounds concerning the distortion of optimal nearest-neighbor quantizers, shrinking cell conditions do appear in more general settings. Na and Neuhoff [7] require shrinking cells in their derivation of Bennett's integral for vector quantizers having convergent point densities and convergent inertial profiles. The same condition appears in recent work [8] on the asymptotic distribution of errors for high-rate vector quantizers.

The r'th power distortion of a quantizer is an average quantity, while its cell diameters measure its worst-case local behavior. In many cases, the connection between quantizer design (which is typically distortion-based) and cell size is not readily apparent. For example, verifying a shrinking cell condition can be problematic when the quantizers under study are designed from finite data sets using iterative or recursive methods that seek to reduce empirical distortion, or when they are defined in terms of secondary quantities such as point densities and inertial profiles.

In Theorem 1 below we establish an asymptotic connection between vanishing r'th power distortion and shrinking cell diameters for quantizers with convex cells. As a consequence, a number of shrinking cell conditions may be easily verified by showing that the quantizers in question have distortion tending to zero. Theorem 1 also plays an important role in the asymptotic analysis [9] of a common greedy growing scheme for tree-structured vector quantizers.

2 Results

A vector quantizer is a map $Q : \mathbb{R}^d \to \mathcal{C}$, where $\mathbb{R}^d$ denotes d-dimensional Euclidean space, and $\mathcal{C} = \{c_1, \ldots, c_m\} \subseteq \mathbb{R}^d$ is a finite set of representative vectors known as the *codebook* of Q. Let P be a fixed probability distribution on $\mathbb{R}^d$. For each $r > 0$, the r'th power distortion of Q with respect to a random vector $X \sim P$ is given by

$$D_r(Q) = E\|Q(X) - X\|^r = \int \|Q(x) - x\|^r dP(x)\,, \tag{1}$$

where $\|\cdot\|$ is the ordinary Euclidean norm on $\mathbb{R}^d$. A sequence of quantizers $Q_1, Q_2, \ldots$ has vanishing r'th power distortion if $D_r(Q_n) \to 0$ as $n \to \infty$.

Every quantizer Q is associated with a finite partition $\{A_1, \ldots, A_m\}$ of $\mathbb{R}^d$, where $A_i = \{x : Q(x) = c_i\}$ is the cell contains those vectors assigned to the i'th codeword. For each vector x the cell of Q containing x is defined by

$$Q[x] = \{u : Q(u) = Q(x)\}\,.$$

The *diameter* of a set $U \subseteq \mathbb{R}^d$ is the greatest distance between any two points of the set, namely

$$\mathrm{diam}(U) = \sup_{u,v \in U} \|u - v\|\,.$$

A sequence of quantizers $Q_1, Q_2, \ldots$ will be said to have *shrinking cells* if for every $\epsilon > 0$

$$P\{x : \mathrm{diam}(Q_n[x]) > \epsilon\} \to 0\,. \tag{2}$$

Equivalently, $\mathrm{diam}(Q_n[X]) \to 0$ in probability when $X \sim P$. Note that $\mathrm{diam}(Q_n[x])$ accounts for all the points in the cell, not just those that lie in the support set of P.

Let $Q_1, Q_2, \ldots$ be vector quantizers such that (i) each cell A_i of Q_n contains its corresponding codeword c_i, and (ii) c_i is no worse a representative than the zero vector in the sense that

$$\int_{A_i} \|x - c_i\|^r dP \leq \int_{A_i} \|x\|^r dP\,.$$

When $r = 2$ both conditions are satisfied if Q_n has convex cells and the representative of each cell is its centroid with respect to P. Under these conditions it is readily verified that shrinking cells imply vanishing distortion.

Proposition 1 *Let $Q_1, Q_2, \ldots$ satisfy conditions (i) and (ii). If $\int \|x\|^r dP < \infty$ and $P\{x : \mathrm{diam}(Q_n[x]) > \epsilon\} \to 0$ for every $\epsilon > 0$, then $D_r(Q_n) \to 0$.* □

Our principle result, stated in Theorem 1 below, is a converse to Proposition 1 that holds if the cells of each quantizer Q_n are convex and P is absolutely continuous with respect to Lebesgue measure. Theorem 1 applies to nearest-neighbor and tree-structured vector quantizers, each of which have convex cells. The proof of the theorem is given in the next section.

Theorem 1 *Let P be an absolutely continuous distribution on $\mathbb{R}^d$ and let $Q_1, Q_2, \ldots$ be vector quantizers having convex cells. If for some $r > 0$ the distortions $D_r(Q_n) \to 0$, then*

$$P\{x : \mathrm{diam}(Q_n[x]) > \epsilon\} \to 0$$

for every $\epsilon > 0$.

Remark: If P fails to be absolutely continuous, then the conclusion of Theorem 1 may not be valid. Nevertheless, the theorem applies to the absolutely continuous part of an arbitrary distribution P.

Remark: Theorem 1 shows that, for quantizers with convex cells, the shrinking cell condition used in [7] is actually implied by their other assumptions.

3 Proof of Theorem 1

The proof of Theorem 1 relies on the following lemma, which is a simple consequence of a geometrical result of Alexander [1]. The lemma provides a connection between the diameter of a cell, and the Lebesgue measure of those points in the cell lying outside a ball with fixed radius but arbitrary center. Let λ denote d-dimensional Lebesgue measure.

Lemma A *Fix $b > 0$ and let $\mathcal{U}$ be the collection of convex sets $U \subseteq \mathbb{R}^d$ such that U has non-empty interior and $\mathrm{diam}(U) \geq b$. For every $\gamma > 0$ there is a number $\delta > 0$ such that*

$$\frac{\lambda(U \cap B(x,\delta))}{\lambda(U)} \leq \gamma$$

for every $U \in \mathcal{U}$ and every $x \in \mathbb{R}^d$. $\Box$

Proof of Theorem 1: It is easiest to establish the contrapositive of the claim: if there exist numbers a, $\epsilon > 0$ such that

$$\limsup_{n\to\infty} P\{x : \mathrm{diam}(Q_n[x]) > a\} > \epsilon\,, \tag{3}$$

then $\limsup_n D_r(Q_n) > 0$ for every $r > 0$. Application of Lemma A requires an inequality analogous to (3) in which diameters are computed within a bounded convex set.

Let C be a bounded rectangle in $\mathbb{R}^d$ so large that $P(C^c) < \epsilon/6$. Our immediate goal is to establish the inequality (6) by contradiction. To this end, suppose that for every $b > 0$,

$$\limsup_{n\to\infty} P\{x : \mathrm{diam}(Q_n[x] \cap C) > b\} \leq \epsilon/3\,. \tag{4}$$

Fix n for the moment and consider the quantizer Q_n. As each cell $Q_n[x]$ is connected,

$$\begin{aligned}
&P\{x : \mathrm{diam}(Q_n[x]) > a\} \\
&\quad\leq\ P\{x : \mathrm{diam}(Q_n[x] \cap C^c) > a/2\,\} + P\{x : \mathrm{diam}(Q_n[x] \cap C) > a/2\} \\
&\quad\leq\ P(C^c) + P\{x \in C : \mathrm{diam}(Q_n[x] \cap C^c) > a/2\,\} + \\
&\qquad\ P\{x : \mathrm{diam}(Q_n[x] \cap C) > a/2\,\}\,.
\end{aligned} \tag{5}$$

If $x \in C$ is such that $\mathrm{diam}(Q_n[x] \cap C^c) > a/2$ then $Q_n[x] \cap C^c \neq \emptyset$, and consequently

$$\inf_{u\in C^c} \|x - u\| \leq \mathrm{diam}(Q_n[x] \cap C)\,.$$

For each $\gamma > 0$ let C_γ be the set of points in C that are close to its boundary:

$$C_\gamma = \left\{x \in C : \inf_{u\in C^c} \|x - u\| < \gamma\right\}\,.$$

Accounting separately for vectors $x \in C_\gamma$ and $x \in C_\gamma^c$ it is evident that for each $\gamma > 0$,

$$P\{x \in C : \mathrm{diam}(Q_n[x] \cap C^c) > a/2\} \ \leq\ P\{x : \mathrm{diam}(Q_n[x] \cap C) > \gamma\} + P(C_\gamma)\,.$$

Combining this last inequality with (4) and (5), it follows that

$$\begin{aligned}\limsup_{n\to\infty} P\{x : \mathrm{diam}(Q_n[x]) > a\} &\leq \frac{\epsilon}{3} + \frac{\epsilon}{3} + \frac{\epsilon}{6} + P(C_\gamma) \\ &= \frac{5\epsilon}{6} + P(C_\gamma)\,.\end{aligned}$$

Since P is absolutely continuous and C is bounded and convex, $P(C_\gamma) \to 0$ as $\gamma \to 0$. When γ is sufficiently small the inequality above contradicts (3), and we conclude that for some $b' > 0$,

$$\limsup_{n\to\infty} P\{x : \mathrm{diam}(Q_n[x] \cap C) > b'\} \geq \epsilon/3 \tag{6}$$

as desired.

Let $\epsilon > 0$ and the bounded rectangle C be as above. Since P is absolutely continuous, there exists a number $\eta > 0$ such that for every measurable $A \subseteq \mathbb{R}^d$, $\lambda(A) \leq \eta$ implies $P(A) \leq \epsilon/12$. By Lemma A there is a number $\delta > 0$ so small that for every vector x and every convex set $U \subseteq \mathbb{R}^d$ with $\mathrm{diam}(U) > b'$,

$$\frac{\lambda(U \cap B(x,\delta))}{\lambda(U)} \leq \frac{\eta}{\lambda(C)}\,. \tag{7}$$

Use of (7) will show that $P\{x : \|x - Q_n(x)\| \geq \delta\}$ is bounded away from zero when n is sufficiently large.

Fix n and let $\{c_1, \ldots, c_N\}$ be the codebook of Q_n. Let $U_j = \{x : Q_n(x) = c_j\}$ be the cell corresponding to the representative vector c_j, and let $Z_n = \{j : \mathrm{diam}(U_j \cap C) > b'\}$ contain the indices of 'large' cells. Define

$$V_n = \bigcup_{j \in Z_n} U_j = \{x : \mathrm{diam}(Q_n[x] \cap C) > b'\}$$

to be the union of the large cells. Consider those vectors $x \in V_n$ that lie close to their representatives. By an obvious upper bound,

$$P\{x \in V_n : \|x - Q_n(x)\| < \delta\} \leq P(C^c) + P\{x \in V_n \cap C : \|x - Q_n(x)\| < \delta\}\,.$$

Summing over the constituent cells of V_n and applying the bound (7) to each cell shows that

$$\begin{aligned}\lambda\{x \in V_n \cap C : \|x - Q_n(x)\| < \delta\} &= \sum_{j \in Z_n} \lambda(U_j \cap C \cap B(c_j, \delta)) \\ &\leq \sum_{j \in Z_n} \eta \frac{\lambda(U_j \cap C)}{\lambda(C)} \\ &\leq \eta\,.\end{aligned}$$

The choice of C and the constant η insures that

$$P\{x \in V_n : \|x - Q_n(x)\| < \delta\} \leq \epsilon/6 + \epsilon/12 = \epsilon/4\,,$$

and consequently

$$\begin{aligned} P\{x : \|x - Q_n(x)\| \geq \delta\} &\geq P\{x \in V_n : \|x - Q_n(x)\| \geq \delta\} \\ &= P(V_n) - P\{x \in V_n : \|x - Q_n(x)\| < \delta\} \\ &\geq P(V_n) - \epsilon/4\,. \end{aligned}$$

By definition of V_n and the relation (6),

$$\limsup_{n\to\infty} P\{x : \|x - Q_n(x)\| \geq \delta\} \geq \epsilon/3 - \epsilon/4 = \epsilon/12\,,$$

and therefore $\limsup_{n\to\infty} D_r(Q_n) > 0$ for every $r > 0$. □

Acknowledgments

We wish to thank Ralph Alexander for bringing the paper [1] to our attention.

References

[1] R. Alexander. A problem about lines and ovals. *American Math. Monthly*, 75(5):482–487, 1968.

[2] J.A. Bucklew and G.L. Wise. Multidimensional asymptotic quantization theory with r'th power distortion measures. *IEEE Trans. Inform. Theory*, 28:239–247, 1982.

[3] S. Cambanis and N. Gerr. A simple class of asymptotically optimal quantizers. *IEEE Trans. Inform. Theory*, 29:664–676, September 1983.

[4] A. Gersho. Asymptotically optimal block quantization. *IEEE Trans. Inform. Theory*, 25:373–380, 1979.

[5] R.M. Gray, K.L. Oehler, K.O. Perlmutter, and R.A. Olshen. Combining tree-structured vector quantization with classification and regression trees. In *Proceedings of the* 27^{th} *Asilomar Conference on Circuits Systems and Computers*, 1993.

[6] G. Lugosi and A.B. Nobel. Consistency of data-driven histogram methods for density estimation and classification. 1996. To appear in Annals of Statistics.

[7] S. Na and D.L. Neuhoff. Bennett's integral for vector quantizers. *IEEE Trans. Inform. Theory*, 41:886–900, 1995.

[8] D.L. Neuhoff. On the asymptotic distribution of the errors in vector quantization. Submitted to IEEE Trans. on Inform. Theory, 1995.

[9] A.B. Nobel. Recursive partitioning to reduce distortion. Technical Report UIUC-BI-95-01, Beckman Institute, University of Illinois, Urbana-Champaign, 1995.

[10] A.B. Nobel. Histogram regression estimation using data-dependent partitions. 1996. To appear in Annals of Statistics.

[11] K.L. Oehler, P.C. Cosman, R.M. Gray, and J. May. Classification using vector quantization. In *Proc. of the Twenty-fifth Asilomar Conference on Signals, Systems and Computers*, pages 439–445, Pacific Grove, Calif., November 1991.

[12] K.L. Oehler and R.M. Gray. Combining image compression and classification using vector quantization. *IEEE Trans. PAMI*, 17:461–473, 1995.

[13] Q. Xie, C.A. Laszlo, and R.K. Ward. Vector quantization for nonparametric classifier design. *IEEE Trans. PAMI*, 15(12):1326–1330, 1993.

[14] Y. Yamada, S. Tazaki, and R.M. Gray. Asymptotic performance of block quantizers with difference distortion measures. *IEEE Trans. Inform. Theory*, IT-26:6–14, 1980.

[15] P. Zador. Asymptotic quantization of continuous random variables. Unpublished memorandum, Bell Laboratories, 1966.

[16] P. Zador. Asymptotic quantization error of continuous signals and quantization dimension. *IEEE Trans. Inform. Theory*, 28:139–149, 1982.