Improved Convergence Rates in Empirical Vector Quantizer Design

András Antos
[1] Informatics Laboratory
Computer and Automation Research Institute
Hungarian Academy of Sciences
Kende u. 13-17, 1111 Budapest, Hungary
e-mail: antos@szit.bme.hu

László Györfi
Department of Computer Science and Information Theory
Budapest Univ. of Technology and Economics
Magyar tudósok krt. 2, 1117 Budapest, Hungary
e-mail: gyorfi@szit.bme.hu

András György[1,2]
[2] Department of Mathematics and Statistics
Queen's University
Kingston, Ontario
Canada, K7L 3N6
e-mail: gya@szit.bme.hu

Abstract — **We consider the rate of convergence of the expected distortion redundancy of empirically optimal vector quantizers. Earlier results show that the mean-squared distortion of an empirically optimal quantizer designed from n independent and identically distributed source samples converges uniformly to the optimum at a rate $O(1/\sqrt{n})$, and that this rate is sharp in the minimax sense. We prove that for any fixed source distribution supported on a given finite set, the convergence rate is $O(1/n)$ (faster than the minimax lower bound), where the corresponding constant depends on the distribution. For more general source distributions, we provide conditions implying a little bit worse $O(\log n/n)$ rate of convergence. In particular, scalar distributions having strictly log-concave densities with bounded support (such as the truncated Gaussian distribution) satisfy these conditions.**

I. Introduction

A d-dimensional N-level *vector quantizer* is a measurable mapping $Q : \mathbb{R}^d \to \mathcal{C}$, where the *codebook* $\mathcal{C} = \{y_1, \ldots, y_N\} \subset \mathbb{R}^d$ is a collection of N distinct d-vectors, called the *code points*. The performance of the quantizer Q when quantizing a random vector $X \in \mathbb{R}^d$ with distribution μ is measured by the *expected (mean-squared) distortion*

$$D(Q) = \mathbf{E}\{\|X - Q(X)\|^2\}.$$

A quantizer Q^* achieving the minimum distortion $D^* = \min_Q D(Q)$ is called *optimal* (thus $D(Q^*) = D^*$).

Assume that the distribution μ is unknown, and a quantizer Q_n is to be designed based on the *training data* $X_1^n = (X_1, \ldots, X_n)$ consisting of n independent and identically distributed copies of X. Note that $Q_n(\cdot) = Q_n(\cdot, X_1, \ldots, X_n)$ is a random function depending on the training data. The goal is to find a quantizer whose performance

$$D(Q_n) = \mathbf{E}\{\|X - Q_n(X)\|^2 | X_1^n\}$$

is close to D^* when quantizing X.

Most design algorithms attempt to find an *empirically optimal vector quantizer* Q_n^* whose distortion is minimal on the training data:

$$D_n(Q_n^*) = \frac{1}{n}\sum_{i=1}^{n} \|X_i - Q_n^*(X_i)\|^2 = \min_Q \frac{1}{n}\sum_{i=1}^{n} \|X_i - Q(X_i)\|^2.$$

To determine the number of training samples necessary to achieve a preassigned level of distortion in quantizing the source X, the finite sample behavior of the *expected distortion redundancy* $\mathbf{E}D(Q_n^*) - D^*$ has to be analyzed.

II. Rate of convergence

Assume that the source distribution μ satisfies the peak power constraint $\mathbf{P}\{\|X\| \le B\} = 1$, that is, μ is concentrated on the sphere $S(B) = \{x \in \mathbb{R}^d : \|x\| \le B\}$. Linder *et al.* [1] proved that, in this case, $\mathbf{E}D(Q_n^*) - D^* \le c/\sqrt{n}$, where the constant c depends on B,N, and d. On the other hand, Bartlett *et al.* [2] showed that the $O(1/\sqrt{n})$ bound on the expected distortion redundancy is tight in the minimax sense: they proved that for any $N \ge 3$ and n large enough, and for any empirically designed N-level quantizer Q_n, there exists a distribution supported in $S(B)$ such that $\mathbf{E}D(Q_n) - D^* \ge \hat{c}/\sqrt{n}$ for some appropriate constant $\hat{c}$ depending on B,N, and d.

However, the minimax lower bound does not describe the behavior of the distortion redundancy for a single fixed source distribution. Next we provide improved individual upper bounds on the expected distortion redundancy. First discrete distributions are considered.

Theorem 1 *Assume that the source distribution μ is concentrated on finitely many atoms. Then*

$$\mathbf{E}D(Q_n^*) - D^* \le c_1/n$$

where the constant c_1 depends on μ and N.

Next we consider more general source distributions. A quantizer Q with codebook $\{y_1, \ldots, y_N\}$ is called *nearest neighbor*, if $\|x - Q(x)\| = \min_i \|x - y_i\|$ for every $x \in \mathbb{R}^d$. For any nearest neighbor quantizer Q, let

$$V(Q) = \min_{\widehat{Q}: D(\widehat{Q}) = D^*} \mathbf{Var}\{\|X - Q(X)\|^2 - \|X - \widehat{Q}(X)\|^2\}$$

where the minimum is taken over all optimal quantizers. Then we have the following result.

Theorem 2 *Assume that $\mathbf{P}\{\|X\| \le B\} = 1$, and let Q_n^* be an empirically optimal quantizer. Assume furthermore that*

$$A = \inf_{Q: D(Q) > D^*} \frac{D(Q) - D^*}{V(Q)} > 0$$

where the infimum is taken over all non-optimal nearest neighbor quantizers whose code points all belong to the sphere $S(B)$. Then

$$\mathbf{E}D(Q_n^*) - D^* \le c_2 \log n/n$$

where the constant c_2 depends on A,B,d, and N.

It can be shown that, for example, scalar source distributions with strictly log-concave densities supported in the interval $[-B, B]$ satisfy the conditions of Theorem 2. Thus, the expected distortion redundancy decreases at a rate $O(\log n/n)$ for such widely used source models as the truncated Gaussian or the uniform distribution.

Acknowledgments

The authors wish to thank Zsolt Bihary and Tamás Linder for useful discussions.

References

[1] T. Linder, G. Lugosi, and K. Zeger, "Rates of convergence in the source coding theorem, in empirical quantizer design, and in universal lossy source coding," *IEEE Trans. Inform. Theory*, vol. 40, pp. 1728–1740, Nov. 1994.

[2] P. Bartlett, T. Linder, and G. Lugosi, "The minimax distortion redundancy in empirical quantizer design," *IEEE Trans. Inform. Theory*, vol. 44, pp. 1802–1813, Sep. 1998.

This research was supported in part by the NATO Science Fellowship of Canada and by a research grant for the Research Group for Informatics and Electronics of the Hungarian Academy of Sciences.