

Rate and Distortion Redundancies for Universal Source Coding with Respect to a Fidelity Criterion

Philip A. Chou[†] and Michelle Effros[‡]

[†] Xerox Palo Alto Research Center, 3333 Coyote Road, Palo Alto, CA 94304

[‡] Information Systems Laboratory, Stanford University, Stanford, CA 94305-4055

Let $\{X_i\} \sim P_\theta$, $\theta \in \Lambda \subseteq \mathbb{R}^K$. Rissanen has shown that there exist universal noiseless codes for $\{X_i\}$ with per-letter rate redundancy as low as $\frac{K \log N}{2N}$, where N is the blocklength and K is the number of source parameters. We derive an analogous result for universal source coding with respect to the squared error fidelity criterion: there exist codes with per-letter rate redundancy as low as $\frac{K \log N}{2N}$ and per-letter distortion (averaged over X^N and θ) at most $D(R)[1 + \frac{K}{N}]$, where $D(R)$ is an average distortion-rate function and K is now the number of parameters in the code.

Let $\{X_i\}$ be a random process over alphabet \mathcal{X} with process measure P_θ , $\theta \in \Lambda$, and let $q^N = \beta^N \circ \alpha^N$ be a length- N block code with encoder $\alpha^N : \mathcal{X}^N \rightarrow \mathcal{C}$ and decoder $\beta^N : \mathcal{C} \rightarrow \mathcal{Y}^N$, where $\mathcal{C} = \{c_1, \dots, c_M\} \subseteq \{0, 1\}^*$ is some binary prefix code and \mathcal{Y} is the reproduction alphabet, typically equal to \mathcal{X} . A universal source code with respect to a fidelity criterion $d(X^N, Y^N) = \sum d(X_i, Y_i)$ is a sequence of block codes $\{q^N\}$ such that for each $\theta \in \Lambda$ there exists a corresponding sequence of points $\{(R_{N,\theta}, D_{N,\theta})\}$ on the graph of the N th order operational distortion-rate function for P_θ for which the per-letter "rate" redundancy

$$\frac{1}{N} \rho_\theta(q^N) = \frac{1}{N} E_\theta |\alpha^N(X^N)| - R_{N,\theta} \quad (1)$$

and the per-letter "distortion" redundancy

$$\frac{1}{N} \delta_\theta(q^N) = \frac{1}{N} E_\theta d(X^N, q^N(X^N)) - D_{N,\theta} \quad (2)$$

each go to zero uniformly in θ (in which case the code is *strongly minimax* universal), pointwise in θ (in which case the code is *weakly minimax* universal), or in expectation with respect to a probability measure on θ (in which case the code is *weighted* universal). In the noiseless case, where $D_{N,\theta} = 0$ and $R_{N,\theta} = H_\theta(X^N)$, Rissanen [1] has shown that when $\Lambda \subset \mathbb{R}^K$ is compact with a non-empty interior, and $\{P_\theta\}$ satisfies certain regularity conditions, there exists a universal code $\{q^N\}$ with per-letter rate redundancy

$$\frac{1}{N} \rho_\theta(q^N) \leq \frac{K \log N}{2N} + o\left(\frac{\log N}{N}\right) \quad (3)$$

for each $\theta \in \Lambda$. (Hence the code is weakly minimax universal.) Rissanen goes on to show that this is also the minimum redundancy achievable by any universal code $\{q^N\}$, for almost all θ (with respect to Lebesgue measure).

We derive a result analogous to (3) for weighted universal source coding with respect to the squared error criterion by analyzing a two-part fixed-rate coding scheme [2], first analyzed by Zeger and Bist [3]. In that scheme, each block $X^N = X^{nk} = (X_1^k, \dots, X_n^k)$ is encoded in two parts. In the first part, a k -dimensional, M -codeword vector quantizer $q^k = \beta^k \circ \alpha^k$ is optimized for the sample distribution of X_1^k, \dots, X_n^k , and then

the M reproduction codewords $\{\beta^k(c)\}$ are themselves encoded using a fixed "universal" vector quantizer optimized for the distribution of the $\beta^k(c)$ s (averaged over all c , X^N , and θ). In the second part of the encoding, X_1^k, \dots, X_n^k are encoded using the quantized code $\hat{q}^k = \hat{\beta}^k \circ \hat{\alpha}^k$.

Let $R = \log M/k$ be the rate in bits per letter of q^k , and let $D_k(R) = ED_{k,\theta}(R)$ be the average k th order operational distortion-rate function $D_{k,\theta}(R)$. We use the high resolution approximation $D_{k,\theta}(R) = C_{k,\theta} e^{-2R}$ and a Lagrangian formulation to determine the optimal bit allocation between q^k and the "universal" quantizer. For this optimal allocation, we show that (with $R_{N,\theta} = R$ in (1) and $D_{N,\theta} = D_{N,\theta}(R)$ in (2)) the per-letter rate and distortion redundancies for the overall code q^{nk} are

$$\frac{1}{N} \rho_\theta(q^{nk}) = \frac{K \log N}{2N} + o\left(\frac{\log N}{N}\right) \quad (4)$$

and

$$\frac{1}{N} E \delta_\theta(q^{nk}) \lesssim D_k(R) \left[1 + \frac{K}{N}\right] - D_N(R), \quad (5)$$

where $K = Mk$ is the total number of parameters in the code q^{nk} . The same results hold in the case where the codewords $\beta(c)$ are scalar quantized. The recent work of Zeger, Bist, and Linder [4] supports these results.

While our rate redundancy result (4) for universal source coding with respect to a fidelity criterion is consistent with Rissanen's result (3) for universal noiseless coding, our distortion redundancy result (5) is consistent with Akaike's result on the expected decrease in log likelihood for empirical maximum likelihood on N samples, with Davisson's result on the expected increase in squared error for empirical linear prediction on N Gaussian samples, and with Pollard's result that the codewords in a quantizer follow a central limit theorem (which implies that the expected increase in squared error is inversely proportional to the number of samples N).

References

- [1] J. Rissanen. Universal coding, information, prediction, and estimation. *IEEE Trans. Information Theory*, 30(4):629-636, July 1984.
- [2] S. Panchanathan and M. Goldberg. Algorithms and architecture for image adaptive vector quantization. In *Proc. Visual Communications and Image Processing*, Cambridge, MA, November 1988. SPIE.
- [3] K. Zeger and A. Bist. Universal adaptive vector quantization using codebook quantization. In *Proc. Int'l Conf. Acoustics, Speech, and Signal Processing*, pages III.381-384, San Francisco, March 1992. IEEE.
- [4] K. Zeger, A. Bist, and T. Linder. Universal source coding with codebook transmission. Preprint, May 1992.