Optimal Strategic Quantizer Design via Dynamic Programming

Anju Anand & Emrah Akyol

Binghamton University

March 6, 2022

This research is supported by NSF grants CCF #1910715 and CAREER CCF #2048042.

Introduction

- Quantization: dividing the source into different regions (using decision levels), with a representation level for each region.
  - Example: $[-1, 1]$ is split as $[-1, t), [t, 1]$
- Decision (encoder) and representative (decoder) levels decided based on some objective.
- Classical quantization setting: encoder and decoder objectives are aligned.
  - Example: minimize $\mathbb{E}\{(x-y)^2\}$, where x is quantized to y
- Strategic quantization: encoder and decoder objectives are misaligned.
  - Example: encoder distortion $\mathbb{E}\{(x^3-y)^2\}$, decoder distortion $\mathbb{E}\{(x-y)^2\}$
- Strategic quantization without cardinality constraint on the message space is the Bayesian persuasion setting.

Strategic quantization problem

- Source $X \in \mathcal{X}$ with probability distribution $\mathbb{P}_X$
- Quantizer $Q : \mathcal{X} \to \mathcal{Z}$ maps X to a message in a discrete set $Z \in \mathcal{Z}$, with cardinality constraint $|\mathcal{Z}| \leq M$.
  - Example: $[-1, t)$ to 1 and $[t, 1]$ to 2
- Decoder generates reconstruction $Y \in \mathcal{Y}$ based on the message Z received, to minimize decoder distortion.
- Misaligned encoder and decoder distortion functions, $\eta_e(X, Y) \neq \eta_d(X, Y)$.
- Encoder designs Q before seeing the realization of X, based only on statistics and the objectives.
- Design of Q involves design of decision levels by the encoder, knowing that the decoder will choose representative levels.
- Distortion functions η_e and η_d, shared prior $\mathbb{P}_X$, quantizer Q are common knowledge
- What is Q at the equilibrium?

Timing of the game

1. Encoder designs a quantizer Q, based on common knowledge (Distortion functions η_e and η_d, shared prior $\mathbb{P}_X$) and announces it to the decoder.
2. Encoder observes a realization of X and generates a message $z \in \mathcal{Z}$ through the announced quantizer $Q : \mathcal{X} \to \mathcal{Z}$ and transmits to the decoder noiselessly.
3. Decoder observes $z \in \mathcal{Z}$, takes action $r \in \mathcal{R}$.

Dynamic programming for non-strategic quantization

- Iterative solutions (Lloyd-Max)[1]
- Dynamic programming was used to avoid the local optima issues in Lloyd-Max algorithm.[2]
- Fixed rate and variable rate constraints were designed.[3]
- Complexity reduction with an assumption on the distortion measure.[4]

---

[1] Gersho and Gray. *Vector quantization and signal compression*. Springer Science & Business Media, 2012.

[2] Bruce. *Optimum quantization*. Tech. rep. MIT Research Laboratory of Electronics, 1965.

[3] Sharma. "Design of absolutely optimal quantizers for a wide class of distortion measures". In: *IEEE Trans. on Inf. Th.* (1978).

[4] Wu. "Quantizer monotonicities and globally optimal scalar quantizer design". In: *IEEE Trans. on Inf. Th.* (1993).

A toy example

- $X \sim U[-1, 1]$
- $M = 3$
- $\eta_e(x, y) = (x^3 - y)^2, \eta_d(x, y) = (x - y)^2$
- Boundaries parameterized as $[-1, r_1), [r_1, r_2), [r_2, 1]$.
- Decoder reconstructions: $y_1 = \frac{-1+r_1}{2}$, $y_2 = \frac{r_1+r_2}{2}$, $y_3 = \frac{r_2+1}{2}$.
- Cost function: $J(r_1, r_2) = \int_{-1}^{r_1}(u^3 - \frac{-1+r_1}{2})^2 du + \int_{r_1}^{r_2}(u^3 - \frac{r_1+r_2}{2})^2 du + \int_{r_2}^{1}(u^3 - \frac{r_2+1}{2})^2 du$
- KKT optimality conditions $\frac{\partial J}{\partial r_1} = \frac{\partial J}{\partial r_2} = 0$
- Only non-degenerate solution: $r_1 = -0.7403, r_2 = 0.7403$

A toy example

- Iteratively enforcing optimality conditions (Lloyd-Max-I) for the encoder and decoder: $r_1 \downarrow -1$ and $r_2 \uparrow 1$
- Does not yield a locally optimal solution since any perturbation of $r_1 = -1, r_2 = 1$ would be preferred by the encoder to $r_1 = -1, r_2 = 1$.
- Dynamic programming based algorithms that yield the globally optimal solutions of this problem.[5][6]

[5] Bruce. *Optimum quantization*. Tech. rep. MIT Research Laboratory of Electronics, 1965.

[6] Sharma. "Design of absolutely optimal quantizers for a wide class of distortion measures". In: *IEEE Trans. on Inf. Th.* (1978).

Dynamic Programming

Assumption: Equilibrium quantizer consists of intervals (convex code-cells).

- Required for dynamic programming derivations presented here.
- Part of the regularity condition in nonstrategic quantization literature.
- The other condition (not assumed here): the representation level lies within the interval considered.

Bellman equations: dividing the interval $[r_0, \alpha]$ into $[r_0, t)$ and $[t, \alpha]$

- $D_m(r_0, \alpha) = \min\limits_{\substack{t \in \mathcal{O} \\ r_0 < t < \alpha}} \left(D_{m-1}(r_0, t) + D_1(t, \alpha) \right)$
- $r_{m-1}(r_0, \alpha) = \arg\min\limits_{\substack{t \in \mathcal{O} \\ r_0 < t < \alpha}} [D_{m-1}(r_0, t) + D_1(t, \alpha)]$

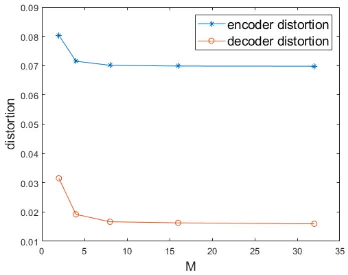
Fixed rate constraint: $|\mathcal{Z}| \leq M$

- Forward pass: $(m-1)^{th}$ decision level for m level quantization of all possible intervals in $\mathcal{X}$.
- Backward pass: optimal decision levels.
- Representative levels found by minimizing decoder's distortion, $y_m = \underset{t \in \mathcal{Y}}{\arg\min} \mathbb{E}(\eta_d(X, Y))$, for each interval.

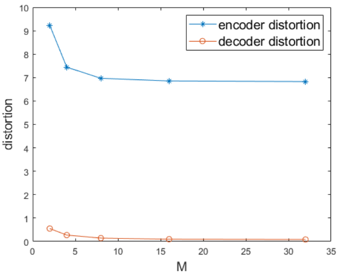
Variable rate constraint: $-\int \log \tau d(\tau) \leq H_0$

- Any quantizer induces a distribution τ over the messages given the prior probability of the source.
- The $1-$level distortion is modified from $D_1(\alpha, \beta)$ to $D_1(\alpha, \beta) + \lambda H_1(\alpha, \beta)$
  - λ - Lagrange parameter
  - $H_1(\alpha, \beta)$ - entropy of quantizing the interval $[\alpha, \beta)$ to 1 level
- The same steps as fixed rate is followed, with iterations over $M = 2, 3, \ldots$ until convergence in $D(\lambda, M) = D_M(x_1, x_K) + \lambda H_M(x_1, x_K)$.
- Assumption: the distortion-rate function, $D(\lambda, M)$, of the optimal quantizer is convex.

Numerical results: Fixed rate



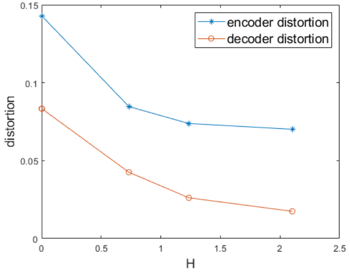
(a) $X \sim U[0,1]$



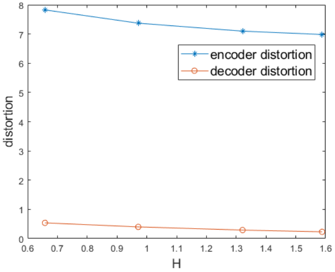
(b) $X \sim \mathbb{N}(0,1)$

Figure: Fixed rate quantization of a uniform and a Gaussian source, for $\eta_e(x,y) = (x^3 - y)^2$ and $\eta_d(x,y) = (x-y)^2$.

Numerical results: Variable rate



(a) $X \sim U[0,1]$



(b) $X \sim \mathbb{N}(0,1)$

Figure: Variable rate quantization of a uniform and a Gaussian source, for $\eta_e(x,y) = (x^3 - y)^2$ and $\eta_d(x,y) = (x-y)^2$.

Results

- Both encoder and decoder distortions monotonically decrease with rate.[7]
- However, unlike their non-strategic counterpart, the distortions stay almost constant as rate increases in the high rate region.
- This is due to the mismatch between objectives of the encoder and the decoder - even if there was no quantization, distortions would not vanish.
- Run time of code for $M = 2, 4, 8, 16, 32$ levels of quantization for the uniform fixed rate case without using complexity reduction versus using complexity reduction: 156.54 seconds versus 126.01 seconds.

[7] Anand and Akyol. *Strategic quantization codes*. 2022. URL: https://github.com/anjuanand98/dpcodes.git.

Discussion

- Early non-strategic quantization literature employed dynamic programming to avoid poor local minima issues in iterative optimization methods such as Lloyd-Max.
- In this paper, we develop dynamic programming algorithms for strategic quantization problem inspired by non-strategic quantization using dynamic programming techniques.
- The usage of dynamic programming here is to resolve issues beyond poor local optima as the iterative solution may not even yield a locally optimal quantizer, as shown with the toy example.

Thank you

Please e-mail any comments/questions to

aanand6@binghamton.edu, eakyol@binghamton.edu