Computational Investigation Reveals Surprising Structure of Optimal Coders for Besov Spaces

Jiashun Jin

Department of Statistics, Stanford University

Methodology

1. Goal: $\approx$ number of bits to ϵ-describe functions in Besov ball $B^{\alpha}_{p,q}(C)$, error measure L^2.
2. Reduce to: $\approx$ minimal number of bits to ϵ-describe sequence in $l^p_n(1)$ ball, error measure l^2.
3. Equivalent to: minimax problem in rate-distortion theory.
4. Minimax problem equivalent to ∞-dimensional linear program.
5. Approximate by: high dimensional linear program.
6. Solved by: Interior point methods.
7. Reaveals structure of optimal codebook.

The Linear Program

1. $(P_{s,t})\ \ \max\ v$: $\int \exp(s(x-y)^2 + t|x|^p)\mu(dy) \geq v$
   where we optimize over probability measure μ on R.
   - Reduction to this problem requires uses of tools of rate-distortion theory and a hypothesis connecting this with Kolmogorov ϵ-entropy. Then hypothesis is proved so far only in case $p = 2$.
   - Interpretation: μ gives the marginal distribution of wavelet coefficients of a given level of wavelet pyramid with s, t chosen.
2. $(\hat{P}_{s,t})\ \ \max\ v$ subject to: $M^T p \geq v\tilde{1},\ p \geq 0,\ 1^T p = 1.$
   where: $p = (p_j), M_{i,j} = \exp(s(j \ominus k)^2 + t|j|^p).$
   $\ominus =$ circular subtraction mod. n.
3. The Dual variables: Associated to solution p of linear program is dual variable $q = (q_i)$,Interpretation: marginal distribution of codebook entries for optimal code for level.

Solving Large-Scale Linear Program

Modern paradigm: primal-dual interior-point (refer to the book: Primal-Dual Interior-Point methods, by Stephen Wright).

- 10-20 iterations.
- Each iteration solves special large-scale linear system.
- We use LSQR, a Krylov subspace method,for our solver.
- Each step of linear system solver involves application of M.
- M is circulant matrix, so fast application can be obtained by FFT.

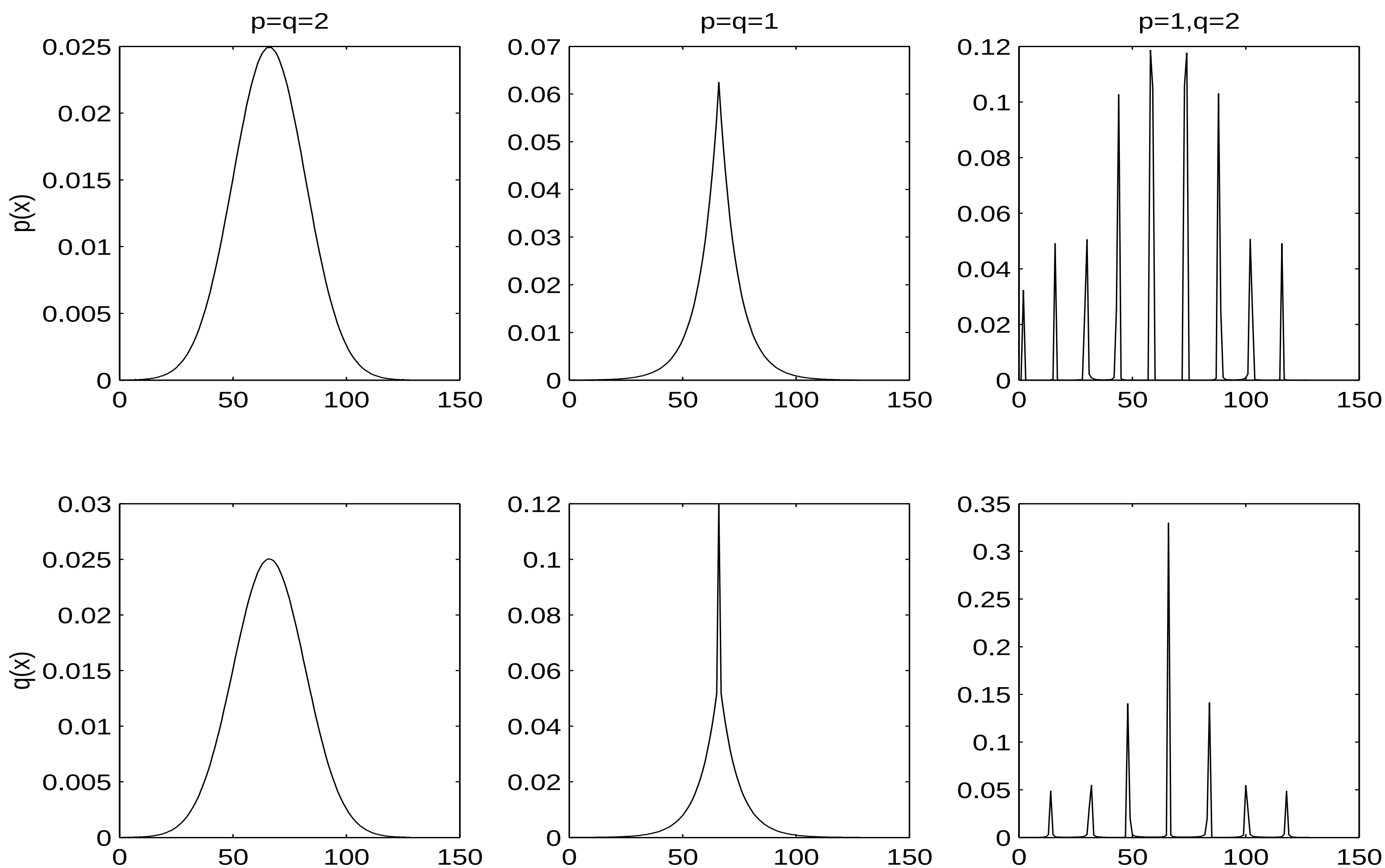

Figure 1: Plots for $p(x)$,$q(x)$ (from top to bottom) in cases $p = q = 2; p = q = 1; p = 1, q = 2$ from left to right.

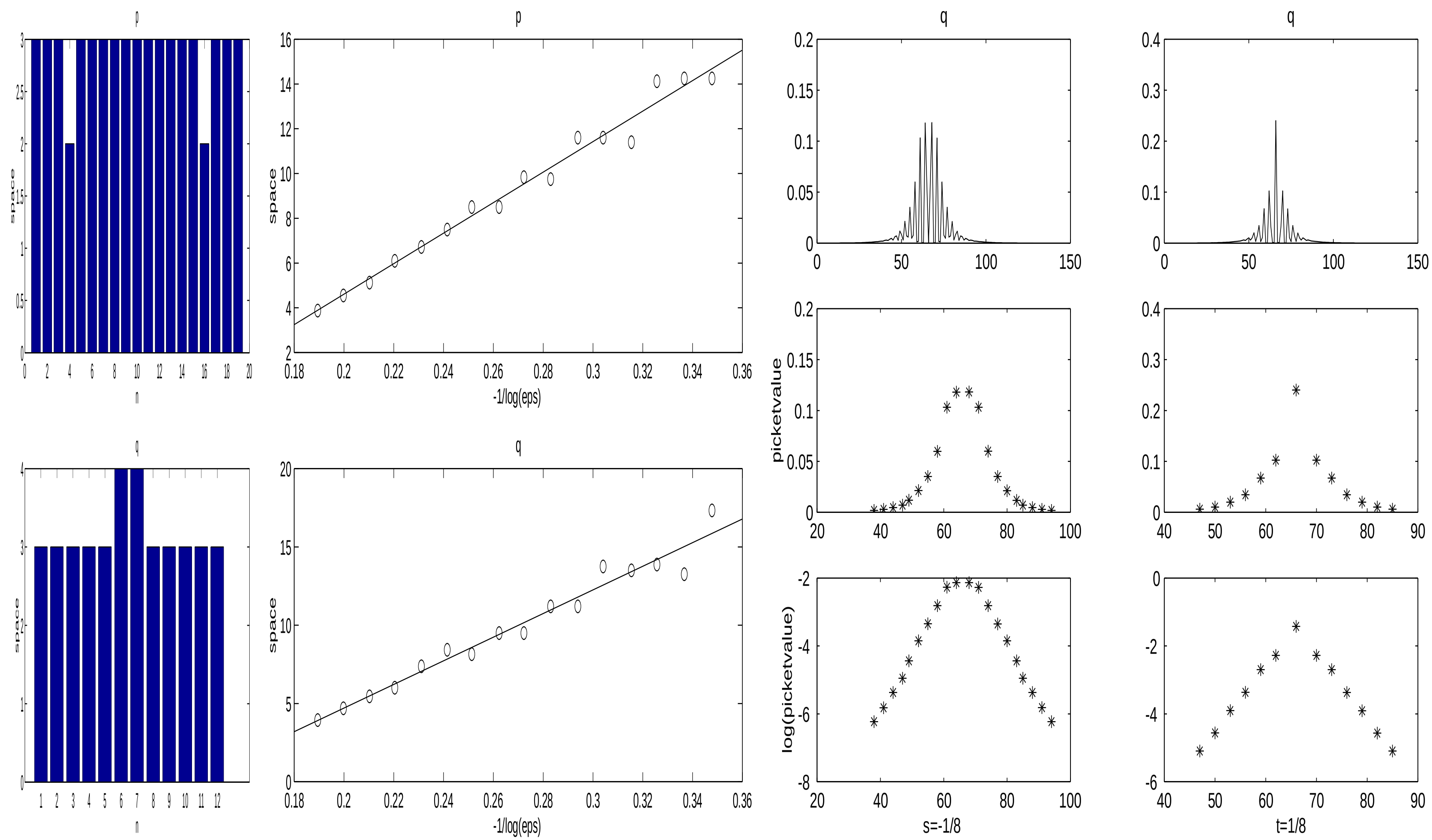

Figure 2: From left to right: barplot shows equal spaced property, plot show space vs $\frac{-1}{\epsilon}$ LINEARLY, and weights decays geometrically.

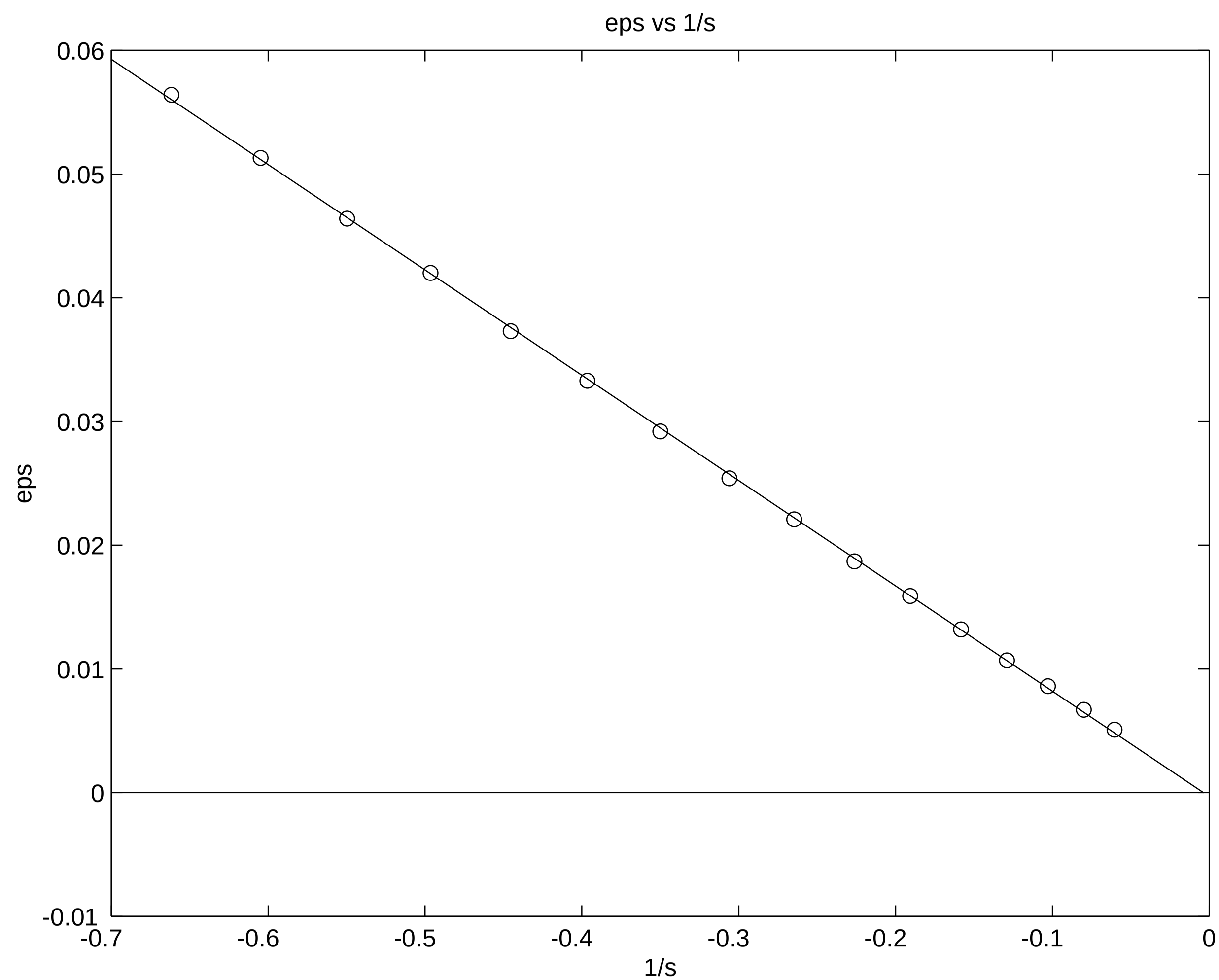

Figure 3: Plots show that $\epsilon = \frac{C}{R'(\epsilon^2)}$,where R(D) is the rate distortion function.

Interpretations and Claims

1. The methodology can find the EXACT constant of $H_\epsilon(\mathcal{F}, \|\cdot\|)$.
2. No previous results on covering l^p-balls by l^q-balls.
3. Properties of Optimal codebook for covering l^p-balls by l^q-balls: a lattice distribution with equi-spaced atoms.
   - Equal spacing of p, q atoms confirms this point.
   - Linearity of Spacing vs $\frac{-1}{\epsilon}$. confirms this point.
   - Geometric decay of p_i and q_i confirms this point.
   - Optimal coding apparently connected to unifrom quantizer.
4. New phenomenon: “crystalline” pattern of optimal covering in high-dimension covering problem.