Failure of Successive Refinement for Symmetric Gaussian Mixtures

James Chow, Student Member, IEEE, and
Toby Berger, Fellow, IEEE
IEEE Transactions on information theory, vol. 43, NO. 1,
Januray 1997

報告人：吳智聖
電信四B
8613010

Introduction to successive refinement

1. When given source and reproduction alphabets Χ and Υ and a source random variable X with distribution p(x), successive refinement from distortion D_1 to distortion D_2 prevails if and only if it is possible to refine an efficient description of the sequence $\{X_k\}$ with average distortion D_1 to a description of $\{X_k\}$ with average distortion $D_2 < D_1$ by providing additional information at rate $\triangle R=R(D_2)-R(D_1)$.

2. The sequence $\{X_k\}$ is said to be successively refinable in general if and only if successive refinement from D_1 to D_2 is achievable for every $D_1 >= D_2$.

3. Furthermore, successive refinement is achievable if the conditional probability distributions $Q(y_1|x)$ and $Q(y_2|x)$ that solve the $R(D_1)$ and $R(D_2)$ problems (i.e., $R(D_i) = I(X; Y_i)$ and $Ed(X; Y_i)$, i =1, 2) form the Markov chain $X - \circ - Y_2 - \circ - Y_1$. That is, successive refinement holds if and only if there exists a probability density $P(y_2|y_1)$ such that

$$P(x|y_1) = \int_{-\infty}^{\infty} P(x|y_2)P(y_2|y_1)dy_2$$

where $P(x|y_1)$ and $P(x|y_2)$ are the backward conditional distributions associated with the solutions to the $R(D_1)$ and $R(D_2)$ problems, respectively.

II. FAILURE OF SUCCESSIVE REFINEMENT

In W. H. R. Equitz and T. M. Cover, "Successive refinement of information" (IEEE Trans. Inform. Theory, vol. 37, no. 2, pp. 269-275, Mar. 1991) successive refinement has been shown to prevail for all finite-alphabet signals with Hamming distortion, for Gaussian signals with squared error distortion, and for Laplacian signals with absolute error distortion. To date, the only example that has been shown not to exhibit successive refinement is Gerrish's problem, in which the input alphabet is discrete. We present an example where successive refinement fails for a continuous-alphabet source.

iid scalar source, symmetric Gaussian mixture

$$p(x) = (1/2)*[N(-\mu, \sigma^2) + N(\mu, \sigma^2)], \quad 0 < \mu < oo$$

$d(x,y) = (x-y)^2$. For $0 <= D <= \sigma^2$

$$q(y) = (1/2)*[N(-\mu, \sigma^2 - D) + N(\mu, \sigma^2 - D)].$$

If $D = \sigma^2$, the output distribution becomes binary, taking on the values $y = \mu$ and $y = -\mu$ each with probability 1/2 .

When $D = D_{max} = \sigma^2 + \mu^2$, the output takes on the value $y = 0$ with probability one.

For $\sigma^2 < D < \sigma^2 + \mu^2$, where the Shannon lower bound is not tight, a result of Rose implies that the reproduction alphabet $\mathcal{Y}$ must be discrete.

If we choose $D_2 = \sigma^2$ and $\sigma^2 < D_1 < \sigma^2 + \mu^2$ i.e.,

$$s_2 = -\frac{1}{2\sigma^2} \quad and \quad -\frac{1}{2\sigma^2} < s_1 < s_{max}$$

then Y_2 is a binary random variable with alphabet $\mathcal{Y}_2 = \{-\mu, \mu\}$ such that $Q_2(y_2) = 1/2$, and the random variable Y_1, which we only take to be discrete, is (2N + 1)-valued with alphabet

$$y_1 = \{\beta_{-N}(s_1), \beta_{-N+1}(s_1), \ldots, \beta_{-1}(s_1), \beta_0(s_1), \beta_1(s_1), \ldots, \beta_{N-1}(s_1), \beta_N(s_1)\}$$

We exploit the symmetry of Y_1 so that $\beta_0 = 0$ with probability γ_0, and $\beta_{-k} = -\beta_k$, k = 1,2,…,N, each with probability γ_k such that

$$\sum_k \gamma_k = 1.$$

We take $P(y_2 \mid y_1)$ as the conditional probability distribution of a discrete, symmetric channel with (2N + 1)-ary input, binary output, and crossover probability

$0 < \alpha_k < 1/2$ given input $y_1 = \beta_k$. The permissible choices of D_1 imply that $Q_1(y_1)$ cannot be the distribution corresponding to that of D_{max}. Hence, we can assume

$Q_1(\beta_1) > 0$ for some $\beta_1 \neq 0$.

To show that successive refinement fails for this example, we must show that equality does not hold in (1).

$$P(x \mid y_i) = \lambda_i(x) p(x) e^{s_i (x - y_i)^2}$$

$$\lambda_i(x)^{-1} = \int_{-\infty}^{\infty} e^{s_i (x - y_i)^2} dQ_i(y_i), \qquad i = 1,2$$

$$\lambda_1(x) p(x) e^{s_1 (x - \beta_1)^2} = \int_{-\infty}^{\infty} \lambda_2(x) p(x) e^{s_2 (x - y_2)^2} P(y_2 \mid \beta_1) dy_2 \qquad (2)$$

$$\frac{\lambda_1(x)}{\lambda_2(x)} = e^{-s_1 (x - \beta_1)^2} (\alpha_1 e^{s_2 (x + \mu)^2} + (1 - \alpha_1) e^{s_2 (x - \mu)^2})$$

$$Q_1(y_1) = \sum_{k=1}^{N} \gamma_k (\delta(y_1 - \beta_k) + \delta(y_1 + \beta_k)) + \gamma_0 \delta(y_1)$$

$$Q_2(y_2) = \frac{1}{2}(\delta(y_2 - \mu) + \delta(y_2 + \mu))$$

$$\frac{\lambda_1(x)}{\lambda_2(x)} = \frac{\frac{1}{2}(e^{s_2(x-\mu)^2} + e^{s_2(x+\mu)^2})}{\sum_{k=1}^{N} \gamma_k (e^{s_1(x-\beta_k)^2} + e^{s_1(x+\beta_k)^2}) + \gamma_0 e^{s_1 x^2}} \quad (4)$$

$$= \frac{\frac{1}{2}(e^{s_2(x^2+\mu^2)}(e^{-2s_2\mu x} + e^{2s_2\mu x}))}{\sum_{k=1}^{N} \gamma_k e^{s_1(x^2+\beta_k{}^2)}(e^{-2s_1\beta_k x} + e^{2s_1\beta_k x}) + \gamma_0 e^{s_1 x^2}} \quad (5)$$

$$\frac{\alpha_1 e^{s_2(x+\mu)^2} + (1-\alpha_1)e^{s_2(x-\mu)^2}}{e^{s_1(x-\beta_1)^2}} = \frac{e^{s_2(x^2+\mu^2)}((1-\alpha_1)e^{-2s_2\mu x} + \alpha_1 e^{2s_2\mu x})}{e^{s_1(x^2+\beta_1{}^2)}e^{-2s_1\beta_1 x}} \quad (6)$$

$$\frac{\frac{1}{2}(e^{-2s_2\mu x} + e^{2s_2\mu x})}{\sum_{k=1}^{N} \gamma_k e^{s_1(x^2+\beta_k{}^2)}(e^{-2s_1\beta_k x} + e^{2s_1\beta_k x}) + \gamma_0 e^{s_1 x^2}} = \frac{(1-\alpha_1)e^{-2s_2\mu x} + \alpha_1 e^{2s_2\mu x}}{e^{s_1(x^2+\beta_1{}^2)}e^{-2s_1\beta_1 x}} \quad \forall x \ni R \quad (7)$$

$$\frac{1}{2}e^{s_1(x^2+\beta_1{}^2)-c_1 x} = \frac{\overline{\alpha_1} e^{-c_2 x} + \alpha_1 e^{c_2 x}}{e^{-c_2 x} + e^{c_2 x}}\left(\sum_{k=1}^{N} \gamma_k e^{s_1(x^2+\beta_k{}^2)}(e^{-c_k x} + e^{c_k x}) + \gamma_0 e^{s_1 x^2}\right) \quad (8)$$

$$-s_1x^2 + c_1x + (\log 2 - s_1\beta_1^2)$$
$$= \log(e^{-c_2x} + e^{c_2x}) - \log(\overline{\alpha}_1 e^{-c_2x} + \alpha_1 e^{c_2x}) - \log[\sum_{k=1}^{N} \gamma_k e^{s_1(x^2+\beta_k^2)}(e^{-c_kx} + e^{c_kx}) + \gamma_0 e^{s_1x^2}] \quad (9)$$

$$-2s_1x + c_1 = c_2 \tanh c_2x + \frac{c_2(\overline{\alpha}_1 e^{-c_2x} - \alpha_1 e^{c_2x})}{\overline{\alpha}_1 e^{-c_2x} + \alpha_1 e^{c_2x}}$$
$$- \frac{\sum_{k=1}^{N} \gamma_k e^{s_1(x^2+\beta_k^2)}[c_k(e^{c_kx} - e^{-c_kx}) + 2s_1x(e^{-c_kx} - e^{c_kx})] + 2\gamma_0 s_1 x e^{s_1x^2}}{\sum_{k=1}^{N} \gamma_k e^{s_1(x^2+\beta_k^2)}(e^{-c_kx} + e^{c_kx}) + \gamma_0 e^{s_1x^2}} \quad (10)$$

When $\alpha_1 = 0$, there are no crossovers in the channel so $Y_1 = Y_2$, and there is no refinement.

When $\alpha_1 =(1/2)$, Y_1 and Y_2 are independent. The Markov relation (1) then gives

$$P(x|y_1) = \int_{-\infty} P(x|y_2)p(y_2)dy_2 = p(x)$$

so X and Y_1 are independent. This corresponds to $D1 = D_{max}$, which is outside the range of interest of $\alpha_2 < D1 < D_{max}$ since s_1 was chosen such that

$$-\frac{1}{2\sigma^2} < s_1 < s_{max} \le 0$$

when $0 < \alpha_1 < (1/2)$. The left-hand side of (10) is a linearly increasing function of x while the corresponding right-hand side is a nonlinear function. Hence, (10) cannot hold continuously in x, and (1) fails to hold. We conclude that successive refinement does not prevail in this example for

$D_2 = \sigma^2$ and $\sigma^2 < D_1 < \sigma^2 + \mu^2$.

III. CONCLUDING REMARKS

We have shown that successive refinement fails to hold for metric Gaussian mixtures.

However, our proof does not explicitly rely on the source density p(x). Rather, knowledge of the source density is utilized only to determine the cardinality $|Y_2|$ of the support of the random variable Y_2 that solves the rate-distortion problem at $D = D_c$, the largest value of D at which the Shannon lower bound is tight.

Furthermore, although symmetry considerations simplified our argument, by no means is symmetry required.

Indeed, we conjecture that successive refinement fails for all continuous i.i.d. scalar sources with squared error distortion when $0 < D_c <= D_2 < D_1 < D_{max}$.

Abstract

We show that information refinement fails for a continuous-alphabet source whose input distribution is the symmetric mixture of normal densities with -μ and μ and common variance σ^2