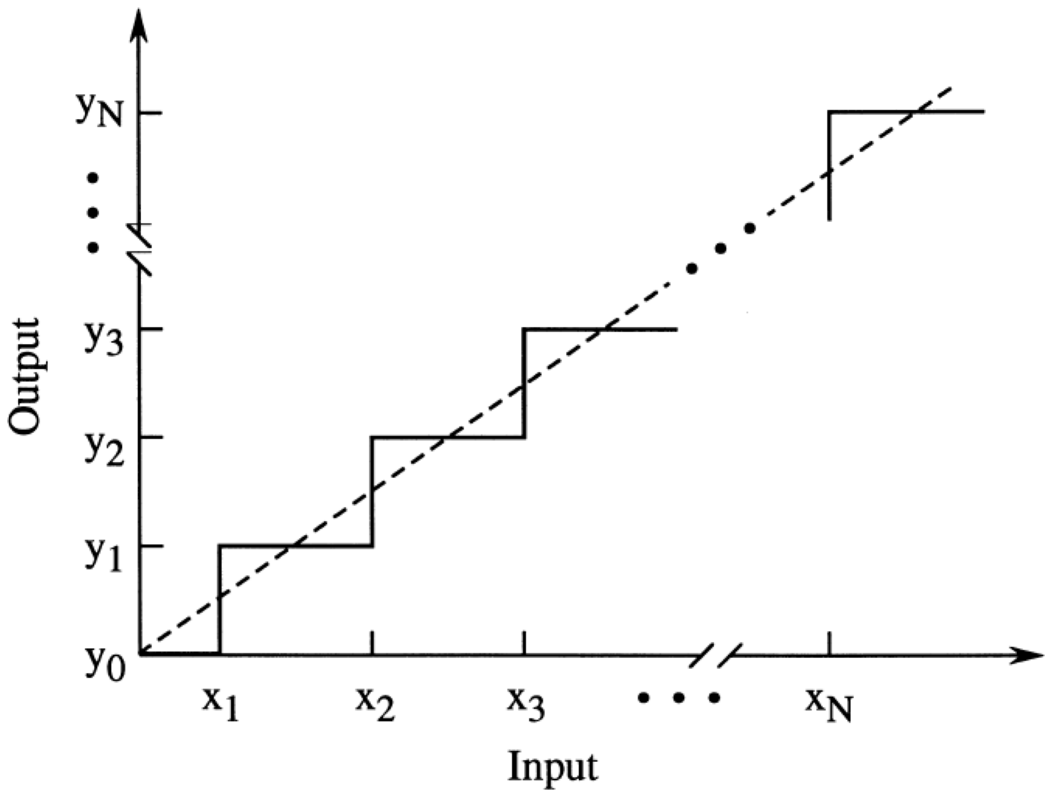
Optimal nonlinear quantizer

General form of quantizer characteristic



$$Q(x) = y_k,\ x_k < x \le x_{k+1}$$

Assume that for any fixed time t, the input signal $x(t)$ is a random variable with probability density function $f_X(x)$.

We wish to choose the N quantizer thresholds $x_k, k = 1, ..., N$ and $N+1$ quantizer output levels $y_k, k = 0, ..., N$ to minimize the expected value of the mean-squared quantization error

$$\phi = \mathrm{E}\left\{\left[X - Q(X)\right]^2\right\}.$$

Here, it is assumed that there are two more threshold levels $x_0 = -\infty$ and $x_{N+1} = \infty$ that are fixed; so that the definition for the quantizer characteristic

$$\left\{ Q(x) = y_k, \quad x_k < x \le x_{k+1} \right\}$$

will be valid for all x on the real line.

We have that

$$\begin{aligned} \phi &= \mathrm{E}\left\{ \left[X - Q(X) \right]^2 \right\} \\ &= \int_{-\infty}^{\infty} \left[x - Q(x) \right]^2 f_X(x) dx \end{aligned}$$

In order to simplify this expression, we substitute the quantizer characteristic into the equation, and then partition the integral into ranges over which the quantizer output is constant

$$\phi = \sum_{k=0}^{N} \int_{x_k}^{x_{k+1}} \left[x - y_k \right]^2 f_X(x) dx$$

This is our fundamental equation for determining the parameters of the optimum nonuniform quantizer.

Determination of optimal quantizer output levels

To determine the optimal value for each y_ℓ for a fixed $0 \le \ell \le N$, we differentiate ϕ with respect to y_ℓ.

$$\frac{\partial \phi}{\partial y_\ell} = \frac{\partial}{\partial y_\ell}\left\{\sum_{k=0}^{N}\int_{x_k}^{x_{k+1}} \left[x - y_k\right]^2 f_X(x)\,dx\right\}$$

$$= \sum_{k=0}^{N}\frac{\partial}{\partial y_\ell}\left\{\int_{x_k}^{x_{k+1}} \left[x - y_k\right]^2 f_X(x)\,dx\right\}$$

Only the $k = \ell$ term under the summation will depend on y_ℓ, and thus be non-zero. So we get

$$\frac{\partial \phi}{\partial y_\ell} = \frac{\partial}{\partial y_\ell}\left\{\int_{x_\ell}^{x_{\ell+1}} \left[x - y_\ell\right]^2 f_X(x)\,dx\right\}$$

$$= \int_{x_\ell}^{x_{\ell+1}} \frac{\partial}{\partial y_\ell}\left\{\left[x - y_\ell\right]^2 f_X(x)\right\}dx$$

$$= \int_{x_\ell}^{x_{\ell+1}} \left\{2\left[x - y_\ell\right]\frac{\partial\left[x - y_\ell\right]}{\partial y_\ell} f_X(x)\right\}dx$$

$$= \int_{x_\ell}^{x_{\ell+1}} \left\{2\left[x - y_\ell\right](-1) f_X(x)\right\}dx$$

Setting $\frac{\partial \phi}{\partial y_\ell} = 0$, canceling out the factors of 2 and -1, and rearranging, we get

$$\int_{x_\ell}^{x_{\ell+1}} y_\ell f_X(x)\,dx = \int_{x_\ell}^{x_{\ell+1}} x f_X(x)\,dx$$

Pulling the constant term y_ℓ out of the integral on the left side of the equal sign, and dividing both sides by the integral $\int_{x_\ell}^{x_{\ell+1}} f_X(x)\,dx$, we get

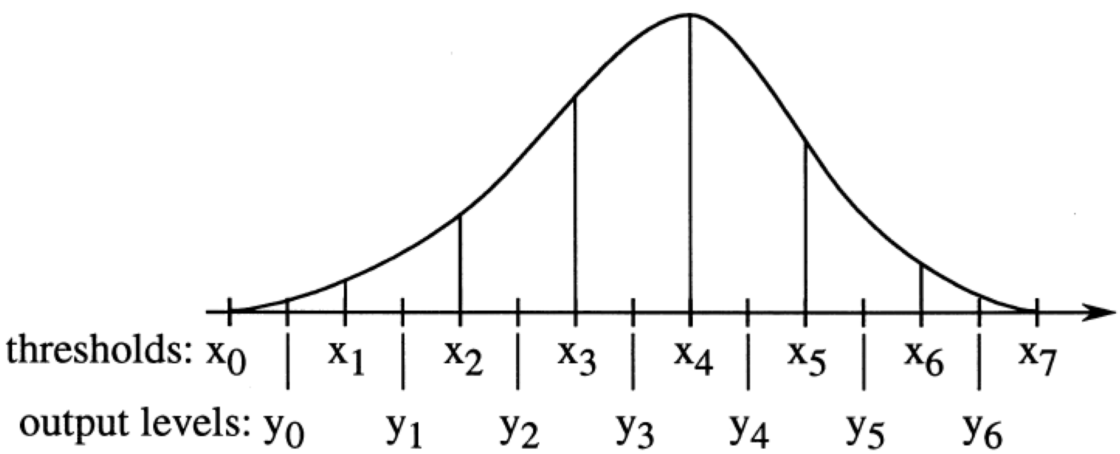
$$y_\ell = \frac{\int_{x_\ell}^{x_{\ell+1}} x f_X(x)\,dx}{\int_{x_\ell}^{x_{\ell+1}} f_X(x)\,dx}$$

$$= \int_{x_\ell}^{x_{\ell+1}} x \left[\frac{f_X(x)}{\int_{x_\ell}^{x_{\ell+1}} f_X(x)\,dx} \right] dx \qquad ,$$

$$= \int_{x_\ell}^{x_{\ell+1}} x f_{X|x_\ell < X \le x_{\ell+1}1}(x | x_\ell < X \le x_{\ell+1})\,dx$$

$$= \mathrm{E}\left\{ X | x_\ell < X \le x_{\ell+1} \right\}$$

which is the mean value for X conditioned on the fact that it lies between x_ℓ and $x_{\ell+1}$.



Determination of optimal quantizer threshold levels

To determine the optimal value for each x_ℓ for a fixed $1 \le \ell \le N$, we similarly differentiate ϕ with respect to x_ℓ.

First, we recall the relationship between the indefinite integral $G(x)$ of an arbitrary function $g(x)$ and the derivative of $G(x)$

$$G(x) = \int_{-\infty}^{x} g(\xi) d\xi$$

$$g(x) = \frac{dG(x)}{dx}$$

To apply this to our particular problem, we define the function

$$\beta(x; y_\ell) = \int_{-\infty}^{x} [\xi - y_\ell]^2 f_X(\xi) d\xi$$

Then we can write

$$\phi = \sum_{k=0}^{N} \int_{x_k}^{x_{k+1}} \left[\xi - y_k\right]^2 f_X(\xi) d\xi$$

$$= \sum_{k=0}^{N} \left[\beta(x_{k+1}; y_k) - \beta(x_k; y_k)\right]$$

Now differentiating with respect to x_ℓ for fixed $1 \le \ell \le N-1$, we obtain

$$\frac{\partial\phi}{\partial x_\ell} = \sum_{k=0}^{N} \left[\frac{d\beta(x_{k+1}; y_k)}{dx_\ell} - \frac{d\beta(x_k; y_k)}{dx_\ell} \right].$$

Note that only two terms under the summation will be non-zero. The term $\frac{d\beta(x_{k+1}; y_k)}{dx_\ell}$ will be nonzero when $k+1=\ell$ or $k=\ell-1$. And the term -$\frac{d\beta(x_k; y_k)}{dx_\ell}$ will be nonzero when $k=\ell$.

Also

$$\frac{d\beta(x; y_\ell)}{dx} = \frac{d}{dx}\left\{ \int_{-\infty}^{x} \left[\xi - y_\ell\right]^2 f_X(\xi) d\xi \right\}$$

$$= \left[x - y_\ell\right]^2 f_X(x)$$

Thus, we have

$$\begin{aligned}\frac{\partial\phi}{\partial x_\ell} &= \frac{\partial}{\partial x_\ell}\left\{\sum_{k=0}^{N}\left[\beta(x_{k+1};y_k)-\beta(x_k;y_k)\right]\right\}\\ &= \frac{d\beta(x_\ell;y_{\ell-1})}{dx_\ell}-\frac{d\beta(x_\ell;y_\ell)}{dx_\ell}\\ &= \left[x_\ell-y_{\ell-1}\right]^2 f_X(x_\ell)-\left[x_\ell-y_\ell\right]^2 f_X(x_\ell)\end{aligned}.$$

Setting the derivative $\frac{\partial\phi}{\partial x_\ell}$ to zero and canceling the common term $f_X(x_\ell)$, we obtain

$$\left[x_\ell-y_\ell\right]^2=\left[x_\ell-y_{\ell-1}\right]^2$$

There are two possible solutions to this equation:

$$\left[x_\ell-y_\ell\right]=\left[x_\ell-y_{\ell-1}\right]$$

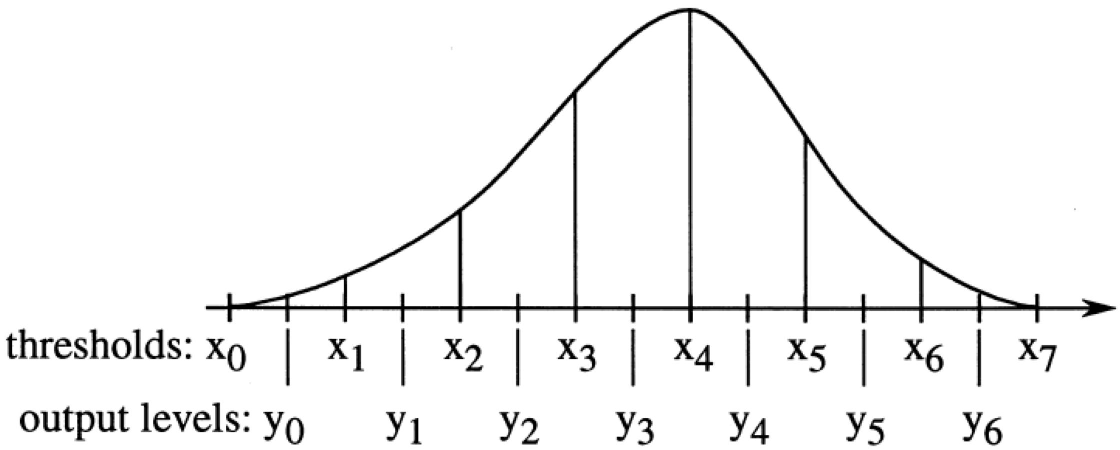
or

$$\left[x_\ell-y_\ell\right]=-\left[x_\ell-y_{\ell-1}\right].$$

The unknown x_ℓ drops out of the first solution; so it is not of interest. The second solution yields

$$x_\ell=\frac{y_\ell+y_{\ell-1}}{2}.$$

This says that each threshold should lie midway between its two neighboring output levels.



Summary of equations for optimal nonuniform quantizer

To summarize, we have two sets of conditions

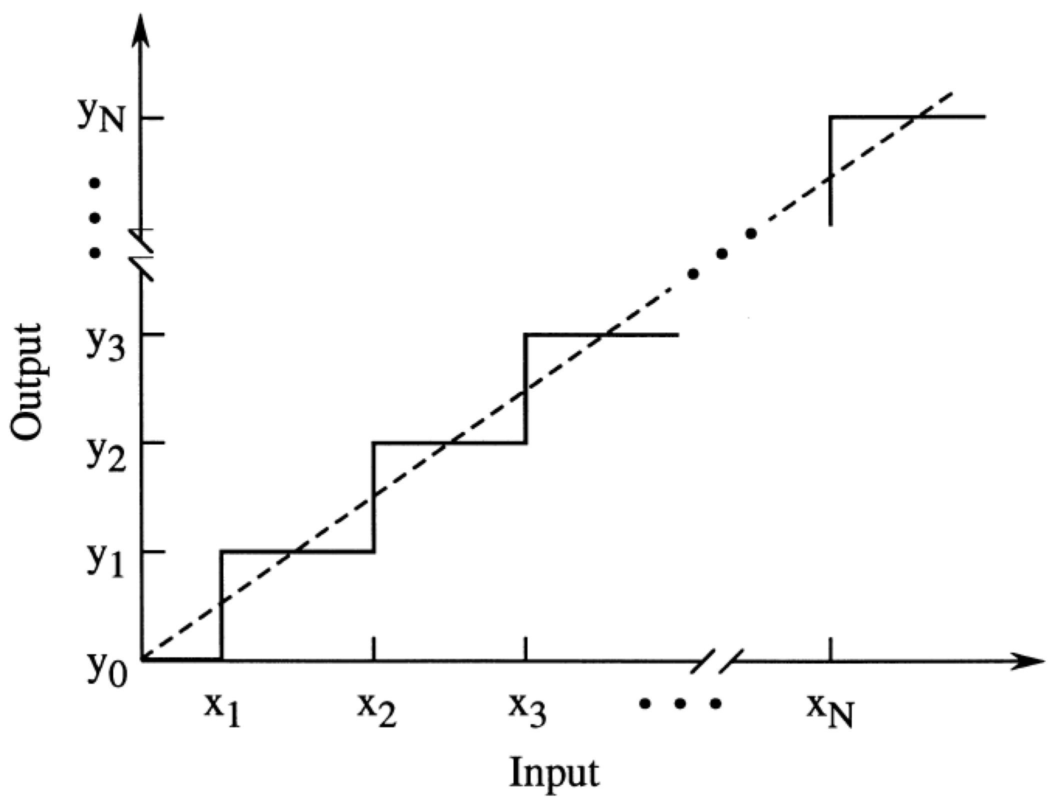
$$y_\ell = \frac{\int_{x_\ell}^{x_{\ell+1}} x f_X(x)\,dx}{\int_{x_\ell}^{x_{\ell+1}} f_X(x)\,dx}, \quad \ell = 0, \ldots, N \tag{1}$$

$$= \mathrm{E}\left\{ X \middle| x_\ell < X \le x_{\ell+1} \right\}$$

and

$$x_\ell = \frac{y_\ell + y_{\ell-1}}{2}, \quad \ell = 1, \ldots, N \tag{2}$$

Note that these equations are coupled. To solve (1), we need to know the thresholds $x_\ell, \ell = 1, \ldots, N$. On the other hand, to solve (2), we need to know the output levels y_ℓ, $\ell = 0, \ldots, N$.



$Q(x) = y_k,\ x_k < x \le x_{k+1}$

How can we deal with this situation?

Lloyd-Max iterative algorithm to determine the parameters of the optimal nonlinear quantizer

We have two coupled sets of conditions (1) and (2) on the previous page that need to be satisfied. The basic idea is to iterate back and forth, alternately updating each set of parameter values one at a time, keeping the other set of parameter values fixed:

1. Initialization: Choose $N+1$ output levels $y_\ell^{(0)}$, $\ell = 0, ..., N$, to be uniformly spaced with some arbitrary constant interval Δ. Set iteration index $k = 0$. Here the superscript 0 for $y_\ell^{(0)}$ denotes the iteration index
2. Update N thresholds: Set

$$x_\ell^{(k)} = \frac{y_\ell^{(k)} + y_{\ell-1}^{(k)}}{2}, \quad \ell = 1, ..., N.$$

3. Update $N+1$ output levels: Set

$$y_\ell^{(k+1)} = \mathrm{E}\left\{ X \middle| x_\ell^{(k)} < X \le x_{\ell+1}^{(k)} \right\}, \quad \ell = 0, ..., N$$

4. Check for completion: (a) The difference

$$\Delta y_\ell^{(k+1)} = y_\ell^{(k+1)} - y_\ell^{(k)}, \quad \ell = 0, ..., N,$$

is sufficiently small (convergence); or (b) the prechosen number of iterations has been completed, i.e. $k = k_{\max}$.
If finished, save current sets

$$x_\ell^{\text{final}} = x_\ell^{(k)}, \quad \ell = 1, ..., N$$
$$y_\ell^{\text{final}} = y_\ell^{(k+1)}, \quad \ell = 0, ..., N$$

Otherwise, increment iteration counter $k \leftarrow k+1$, and return to Step 2 above.

Comments

1. The Lloyd-Max algorithm was first published by Max. But later it was discovered that Lloyd had earlier proposed the same idea in an unpublished memorandum.
2. The algorithm is guaranteed to converge. But it might converge to a local minimum.
3. To implement the algorithm, the expressions (1) and (2) on p. 8 can be evaluated analytically based on the given probability density function $f_X(x)$ for the random variable X, or if this is not possible, they can be evaluated by generating sample outcomes of the random variable, i.e. a histogram, and then performing numerical computations.
4. The Lloyd-Max algorithm can be generalized to vector-valued data. The resulting optimal vector quantizer is known as the Linde-Buzo-Gray algorithm. One application where it has been used is to determine the optimal set of colors $\left\{(R,G,B)_k, k=1,...,N\right\}$ and the mapping of an arbitrary color (R,G,B) to the closest color in this set. This is the topic of the latter part of Module 2.3.4 in the legacy course notes.
5. The Linde-Buzo-Gray algorithm is essentially the same as the well-known K-Means algorithm that is used in statstics for clustering multidimensional datasets.
6. K-Means is still widely used, and today would be regarded as an early example of a machine learning algorithm.
7. Matlab provides built-in support for these algorithms.

8. The idea of iterating back and forth, optimizing over one set of parameters, while holding the other set of parameters constant, and then switching the roles of the two sets is very powerful. It is used in many different settings. One example is the Expectation-Maximization (E-M) algorithm that is widely used in signal reconstruction and machine learning.