CT216 Introduction to Communication Systems
Lecture 5: Quantization of Sampled Signal

Yash M. Vasavada

Professor, DA-IICT, Gandhinagar

6th March 2024

Overview of Today's Talk

Overview of Today's Talk

Overview of Today's Talk

Overview of Today's Talk

Overview of Today's Talk

Overview of Today's Talk

Overview of Today's Talk

Overview of Today's Talk

Overview of Today's Talk

Overview of Today's Talk

Overview of Today's Talk

Overview of Today's Talk

Overview of Today's Talk

Overview of Today's Talk

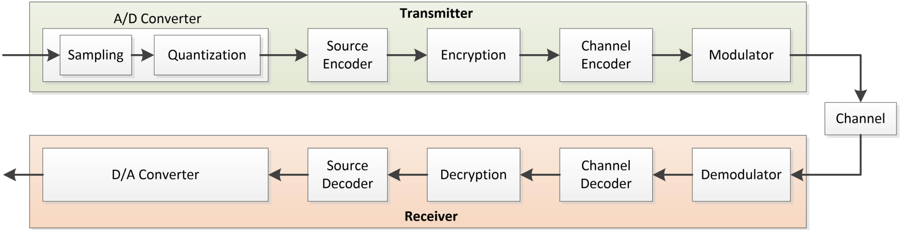
Overview of Today's Talk

Overview of Today's Talk

Digital Communication Transceiver

Block Diagram

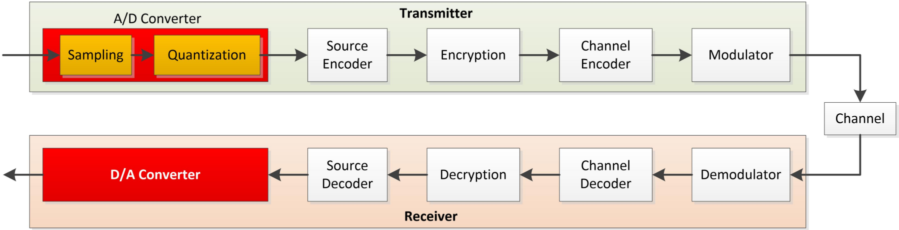
- In CT-542, we have studied this block diagram model of a digital communication transceiver



Digital Communication Transceiver

Block Diagram

- We will now be looking at the process of quantization of an analog information source



Digital Representations of Analog Signals

- Analog signals (video, voice, image, etc.) are continuous in nature
- Sampling theorem says that if the analog signal is band limited, it can be exactly recovered by its time domain samples taken sufficiently close together
  - → Sampling analog signals makes them discrete in time
  - → For band-limited signal, sampling introduces no distortion
- Quantization of sampled analog signals makes the samples discrete in amplitude
  - → Quantization introduces some distortion. There is a trade off between the bandwidth requirement and distortion

Sampling Theorem

- Let $x(t)$ be a band-limited signal with Fourier Transform $X(f)$ which is zero if $|f| > B$
- Time domain signal $x(t)$ can be perfectly reconstructed from its uniformly spaced samples provided these samples are taken at a rate $R > 2B$. Here, $2B$ is called the Nyquist Rate
- If time-domain samples are collected at a rate less than $2B$, aliasing occurs and it is not possible to perfectly reconstruct the analog signal from its samples
- Proof of sampling theorem is based on the following two theorems:
  - → Fourier Transform of the time domain impulse train whose impulses are separated by $\dfrac{1}{2B}$ seconds is the frequency domain impulse train whose consecutive samples are spaced $2B$ Hertz apart
  - → Multiplication of two time domain signals is same as convolution of these two signals in frequency domain

Quantization

Analog Samples → Digital Samples

This is to be covered in **today's** lecture.

- Source coding compresses the digital data with a scheme that provides perfect recovery. No loss of information.
- Continuous-valued signals: infinite number of bits required to represent each signal sample so that it can be reproduced exactly
- Quantization:
  - → approximates continuous-valued samples with finite number of bits
  - → Distortion cannot be avoided
  - → Source coding can be performed on quantized signal samples

Quantization

Analog Samples $\rightarrow$ Digital Samples

- Initial part of this topic will be from Proakis' book Chapter 3.
- Following are some very useful references for the design of Vector Quantizers:
  - $\rightarrow$ R. M. Gray, "Vector Quantization," IEEE ASSP Magazine, April 1984, pp. 4-29
  - $\rightarrow$ A. Gersho and R. M. Gray, "'Vector Quantization and Signal Compression," Kluwer Academic Publications, Boston, 1992.
  - Duda and Hart book on Pattern Classification is the generic reference

Notations

- X: a random variable representing a sample of continuous-valued signal
- $\tilde{X} = f_Q(X)$ is the quantized value of X
- Quantization scheme: if $x_{k-1} \leq x < x_k \Rightarrow \tilde{X} = f_Q(X) = \tilde{x}_k$
  - $\rightarrow$ $\{x_0, x_1, \ldots, x_L\}$: endpoints of L quantization regions, where $x_0 = -\infty$ and $x_L = \infty$
  - $\rightarrow$ L : Number of levels of quantizer
  - $\rightarrow$ Quantizer output or levels: $\tilde{X} \in \{\tilde{x}_1, \tilde{x}_2, \ldots, \tilde{x}_L\}$
    - ▷ For most purposes, a quantizer can be represented by the above list of quantization levels

An Example

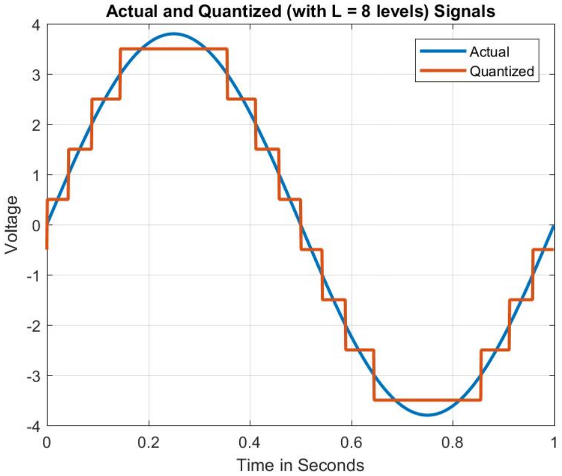
of a Quantizer

An example quantizer with $L = 8$.

k	x_{k-1}	x_k	$\tilde{x}_k$	**Output Bits**
1	$-\infty$	-3	-3.5	000
2	-3	-2	-2.5	001
3	-2	-1	-1.5	010
4	-1	0	-0.5	011
5	0	1	0.5	100
6	1	2	1.5	101
7	2	3	2.5	110
8	3	∞	3.5	111

An Example

Quantization of a Sinusoidal Signal



Notations (Contd...)

Quantization

- Rate R of a quantizer is the number of bits required to represent a sample
- $R = \log_2 L$ bits/sample
- In the previous example, $R = 3$ bits / sample
- *Uniform quantizer* is the one in which $\tilde{x}_k - \tilde{x}_{k-1} = \Delta, \forall k \in 1, \ldots, L-1$
- In a *non-uniform quantizer*, the quantization regions are of unequal widths
  - → when would a non-uniform quantizer be preferred over a uniform quantizer?
- Uniform versus Non-uniform quantization is sometimes also called linear versus nonlinear quantization coding

Distortion

due to Quantization

- Every quantizer introduces distortion into the signal.
- Let us introduce the notion of a distortion function $d(x, \tilde{x})$
- $d(x, \tilde{x})$ between two numbers x and $\tilde{x}$ can be any function, but (intuitively) it should at least be a non-decreasing function of $|x - \tilde{x}|$
- We want to minimize the average distortion D defined as follows:

$$
\begin{aligned}
D = E\left[d\left(X, \tilde{X}\right)\right] &= \int_{-\infty}^{\infty} d(x, \tilde{x}) p(x)\, dx \\
&= \sum_{k=1}^{L} \int_{x_{k-1}}^{x_k} d(x, \tilde{x}) p(x)\, dx
\end{aligned}
$$

Mean Squared Error or MSE

in the Quantization Process

- A common choice for the distortion function: $d(x, \tilde{x}) = (x - \tilde{x})^2$
- $D = MSE = E\left[\left(X - \tilde{X}\right)^2\right] = \sum_{k=1}^{L} \int_{x_{k-1}}^{x_k} (x - \tilde{x})^2 \, p(x) \, dx$
- An interpretation of MSE: $\tilde{X} = f_Q(X) = x + \tilde{n}$, where $\tilde{n} = x - \tilde{x}$ is the noise introduced by the quantization operation. In this case, MSE $= E\left[\tilde{n}^2\right]$ may be interpreted as the power of the quantization noise
- $\left(\frac{P_S}{P_N}\right) = \frac{\text{Signal Power}}{\text{Noise Power}} = \frac{E\left[X^2\right]}{D}$
- When quantization noise $\tilde{n}$ is uniformly distributed, quantization-induced SNR is given approximately as $6 \times R$ dB.

MSE

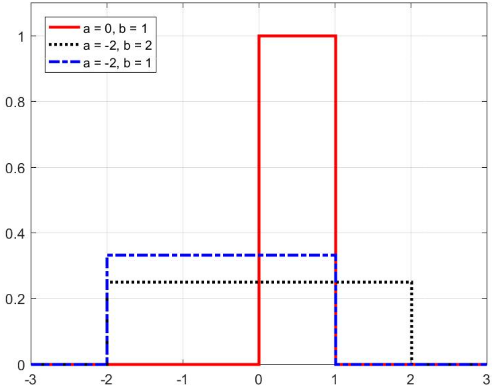
An Example

- Let $L = 8$, $\{\tilde{x}_1 = -3.5, \tilde{x}_2 = -2.5, \ldots, \tilde{x}_L = 3.5\}$
- Let $p(x) = \begin{cases} \frac{1}{8}, & -4 \leq x < 4 \\ 0, & \text{otherwise} \end{cases}$
- $E\left[X^2\right] =?$
- $\text{MSE} = D =?$

Recall:

Uniform PDF

- $p(x) = \begin{cases} \frac{1}{b-a}, & a \leq x \leq b \\ 0, & \text{else} \end{cases}$



Recall:

Uniform PDF

- Mean: $m_x = \int_a^b x\, p(x)\, dx = \frac{1}{b-a}\int_a^b x\, dx = \frac{a+b}{2}$
- Variance: $\sigma_x^2 = \int_a^b (x - m_x)^2\, p(x)\, dx = \frac{(b-a)^2}{12}$
- Probability:

$$P(a_1 \le x < b_1) = \int_{a_1}^{b_1} p(x)\, dx = \frac{b_1 - a_1}{b - a}, \quad a < a_1, b_1 < b$$

MSE

An Example

- $E\left[X^2\right] = \frac{(b-a)^2}{12} = \frac{64}{12}$, since $b = -a = 4$
- Distortion is evaluated as follows:

$$
\begin{aligned}
D &= \sum_{k=1}^{L} \int_{x_{k-1}}^{x_k} (x - \tilde{x})^2 \, p(x) \, dx \\
&= \sum_{k=1}^{8} \int_{-5+k}^{-4+k} (x - (-4.5 + k))^2 \, (1/8) \, dx \\
&= \frac{1}{12}
\end{aligned}
$$

- SNR: $\left(\frac{P_S}{P_N}\right) = \frac{64/12}{1/12} = 64$ (which, in the decibel scale, equals 18 dB)

Quantization

A Design Goal

- Goal of the quantizer design is to
  - $\rightarrow$ minimize D for a given rate R, or
  - $\rightarrow$ minimize R for a given distortion D
- **A Fundamental Theorem of Information Theory:** for a given distortion measure $d(x, \tilde{x})$, there exists a Rate Distortion Function $R(D)$ that provides the lower bound. Achievable rate R given a distortion D cannot be less than the Rate Distortion Function $R(D)$.

Quantization

Rate Distortion Function

- Let X be a zero mean Gaussian variable with a PDF of

$$p(x) = \frac{1}{\sqrt{2\pi\sigma_x^2}} \exp\left(-\frac{x^2}{2\sigma_x^2}\right)$$

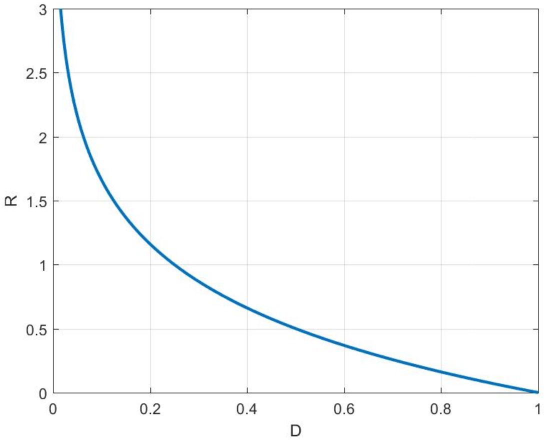
- For this PDF, the Rate Distortion Function takes the following form:

$$R(D) = \begin{cases} \frac{1}{2}\log_2\left(\frac{\sigma_x^2}{D}\right), & 0 \leq D \leq \sigma_x^2 \\ 0, & \sigma_x^2 < D \end{cases}$$

- Alternatively, $D(R) = \sigma_x^2 \times 2^{-2R}$
- Above two mathematical functions are shown on the next two slides for a case when $\sigma_x^2 = 1$
- Suppose a quantizer is designed for this Gaussian source for which $\tilde{x}_1 = -1.494, \tilde{x}_2 = -0.498, \tilde{x}_3 = 0.498, \tilde{x}_4 = 1.494$
  - $\rightarrow$ Numerical integration shows that $D = 0.1188$. What is the theoretical lower bound?

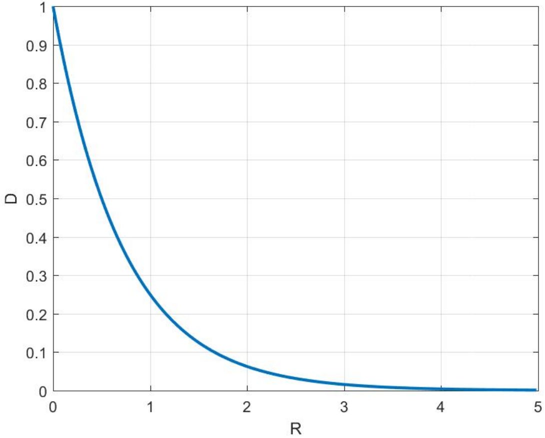
Quantization

Rate Distortion Function



Quantization

Rate Distortion Function



A general strategy

- Consider the distortion measure: $D = \int_{-\infty}^{\infty} (x - \tilde{x})^2 \, p(x) dx$
- Objective: minimize D given a certain rate R
- A general strategy:
  - $\rightarrow$ Make $(x - \tilde{x})^2$ where $p(x)$ is large
  - $\rightarrow$ Tolerate larger $(x - \tilde{x})^2$ where $p(x)$ is small
- i.e., concentrate quantization levels in the regions of large PDF.
- Question: how to do so in an algorithmic manner?

Two Questions

Question 1

- We are given a set of quantization levels $\{\tilde{x}_1, \tilde{x}_2, \ldots, \tilde{x}_L\}$.
- How to select the boundaries of quantization regions $\{x_0, x_1, \ldots, x_L\}$ so as to minimize D?
- Answer: quantize an analog sample x to the nearest quantization level $\tilde{x}$ so as to minimize the squared error $(x - \tilde{x})^2$.
- This is achieved by setting $x_k = \dfrac{\tilde{x}_k + \tilde{x}_{k+1}}{2}, 1 \leq k \leq L - 1$.

Two Questions

Question 2

- We are given a set of quantization boundaries $\{x_0, x_1, \ldots, x_L\}$.
- How to select the quantization levels $\{\tilde{x}_1, \tilde{x}_2, \ldots, \tilde{x}_L\}$ so as to minimize D?
- Answer: differentiate D with respect to $\tilde{x}_k$ and set the derivative to zero.

Two Questions

Question 2

$$\frac{dD}{d\tilde{x}_k} = \frac{1}{d\tilde{x}_k}\sum_{i=1}^{L}\int_{x_{i-1}}^{x_i}(x-\tilde{x}_i)^2\,p(x)dx = 0$$

$$\Rightarrow \frac{1}{d\tilde{x}_k}\int_{x_{k-1}}^{x_k}(x-\tilde{x}_k)^2\,p(x)dx = 0$$

$$\Rightarrow -2\int_{x_{k-1}}^{x_k}(x-\tilde{x}_k)\,p(x)dx = 0$$

$$\Rightarrow \tilde{x}_k\int_{x_{k-1}}^{x_k}p(x)dx = \int_{x_{k-1}}^{x_k}xp(x)dx$$

$$\Rightarrow \tilde{x}_k = \frac{\int_{x_{k-1}}^{x_k}xp(x)dx}{\int_{x_{k-1}}^{x_k}p(x)dx}$$

- Thus, $\tilde{x}_k$ should be located at the *centroid* of the k^{th} quantization region.

Lloyd-Max Algorithm

Summary

1. Choose an initial set of quantization levels: $\{\tilde{x}_1, \tilde{x}_2, \ldots, \tilde{x}_L\}$
2. Update the quantization level boundaries:
   $x_k = \dfrac{\tilde{x}_k + \tilde{x}_{k+1}}{2}, 1 \leq k \leq L - 1$
3. Update the quantization levels: $\tilde{x}_k = \dfrac{\displaystyle\int_{x_{k-1}}^{x_k} xp(x)dx}{\displaystyle\int_{x_{k-1}}^{x_k} p(x)dx}$
4. (optional) evaluate the resulting distortion D_k
5. Repeat steps 2 and 3 until convergence occurs, i.e., either when $D_k - D_{k-1} < \epsilon$ or when $|\tilde{x}_k - \tilde{x}_{k-1}| < \epsilon, \forall k$.

Lloyd-Max Algorithm

Exponential PDF

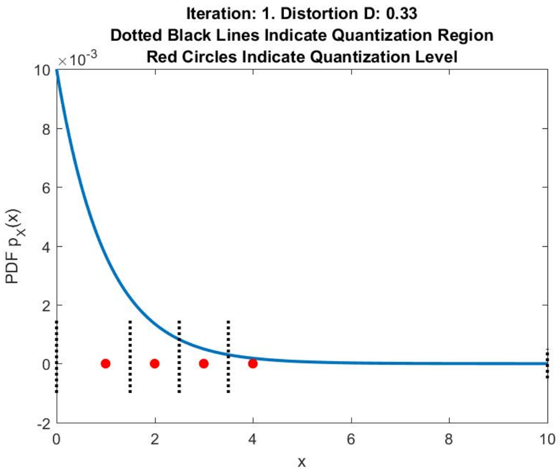
- Lloyd-Max algorithm is applied to the Exponential PDF given as

$$p(x) = \begin{cases} \exp(-x), & x \geq 0 \\ 0, & x < 0 \end{cases}$$

- $L = 4$, with initial quantization levels $\{\tilde{x}_1, \tilde{x}_2, \ldots, \tilde{x}_L\}$ set to $\{1, 2, 3, 4\}$.
- Initial quantization boundaries are found from $x_k = \dfrac{\tilde{x}_k + \tilde{x}_{k+1}}{2}, 1 \leq k \leq L - 1$ as $\{0, 1.5, 2.5, 3.5, \infty\}$.
- Subsequent iterations of Lloyd-Max algorithm are depicted in the following slides.

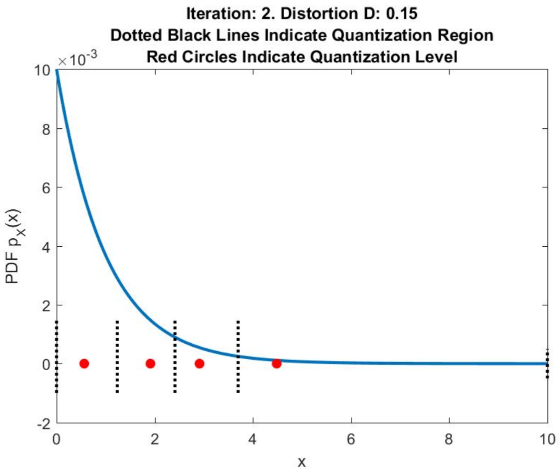
Lloyd-Max Algorithm

Exponential PDF: Initialization (Iteration 1)



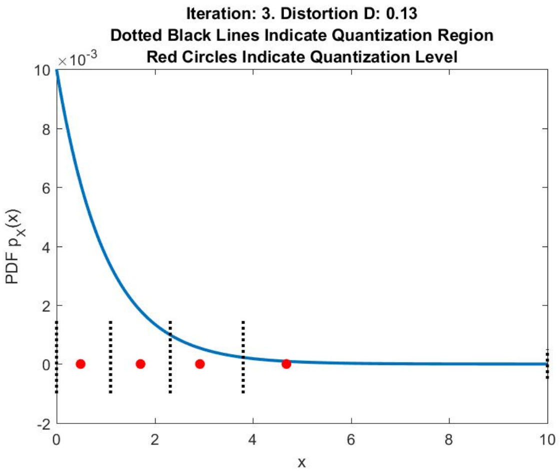
Lloyd-Max Algorithm

Exponential PDF: Iteration 2



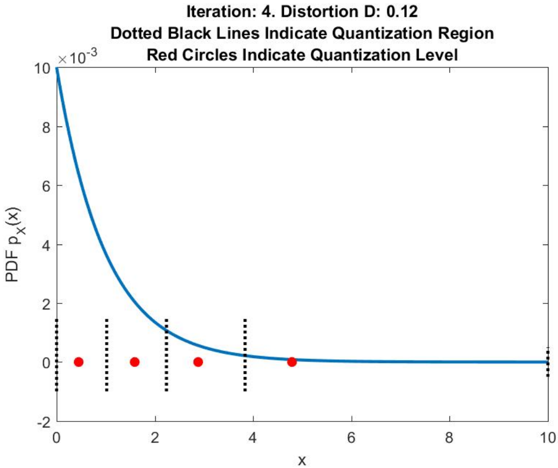
Lloyd-Max Algorithm

Exponential PDF: Iteration 3



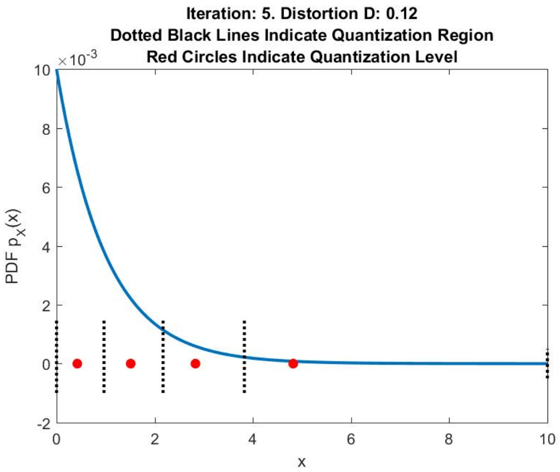
Lloyd-Max Algorithm

Exponential PDF: Iteration 4



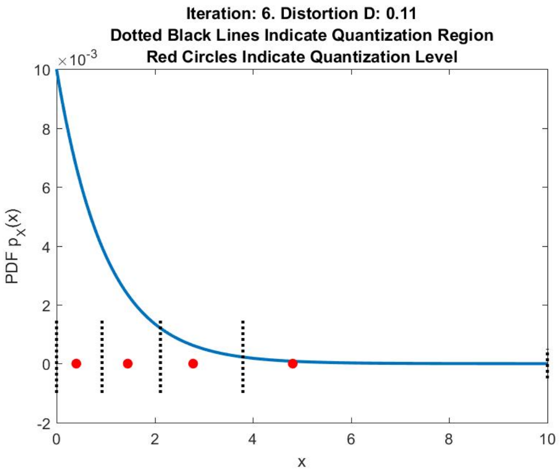
Lloyd-Max Algorithm

Exponential PDF: Iteration 5



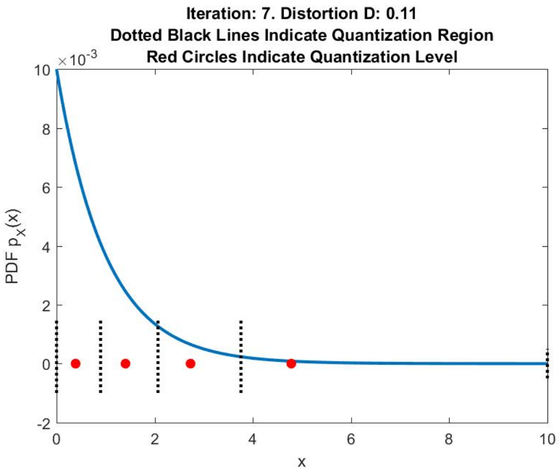
Lloyd-Max Algorithm

Exponential PDF: Iteration 6



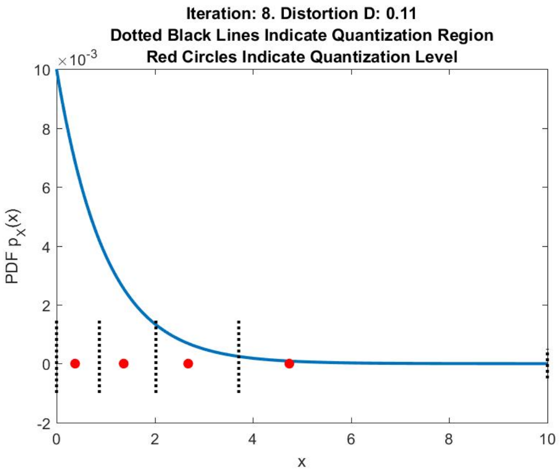
Lloyd-Max Algorithm

Exponential PDF: Iteration 7



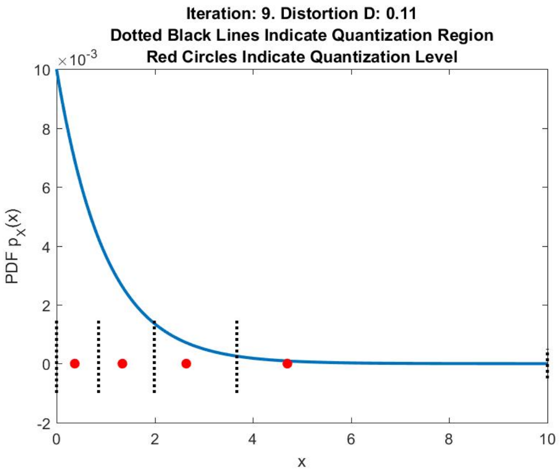
Lloyd-Max Algorithm

Exponential PDF: Iteration 8



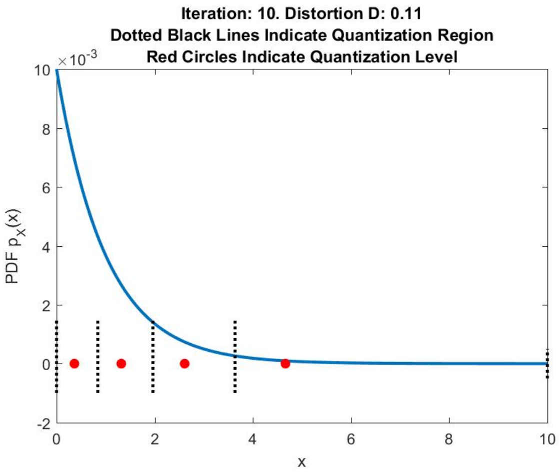
Lloyd-Max Algorithm

Exponential PDF: Iteration 9



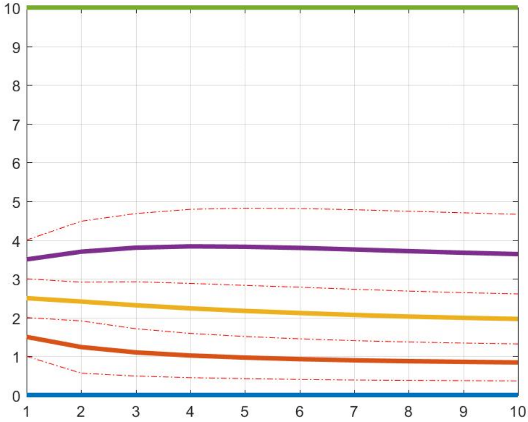
Lloyd-Max Algorithm

Exponential PDF: Iteration 10



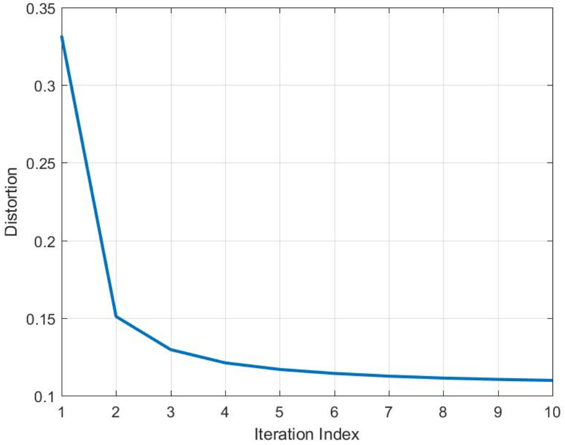
Lloyd-Max Algorithm

Exponential PDF: Algorithm Convergence



Lloyd-Max Algorithm

Exponential PDF: Algorithm Convergence



Lloyd-Max Algorithm

Gaussian PDF

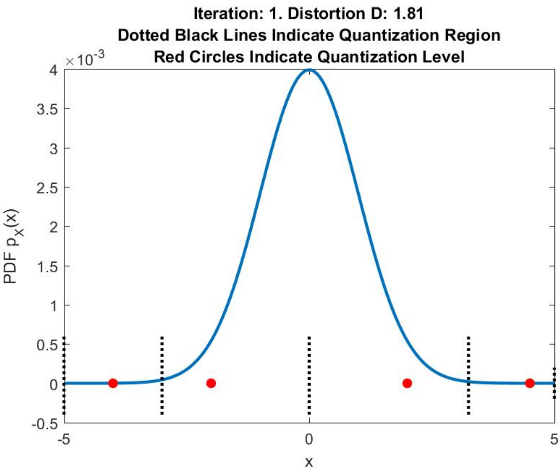
- Lloyd-Max algorithm is applied to the Gaussian PDF given as

$$p(x) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2}\right)$$

- $L = 4$, with initial quantization levels $\{\tilde{x}_1, \tilde{x}_2, \ldots, \tilde{x}_L\}$ set to $\{-4, -2, 2, 4\}$.
- Initial quantization boundaries are found from $x_k = \dfrac{\tilde{x}_k + \tilde{x}_{k+1}}{2}, 1 \leq k \leq L-1$ as $\{-\infty, -3, 0, 3, \infty\}$.
- Subsequent iterations of Lloyd-Max algorithm are depicted in the following slides.

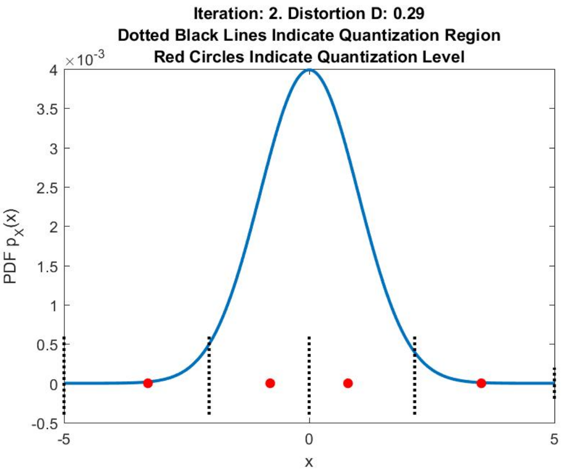
Lloyd-Max Algorithm

Gaussian PDF: Initialization (Iteration 1)



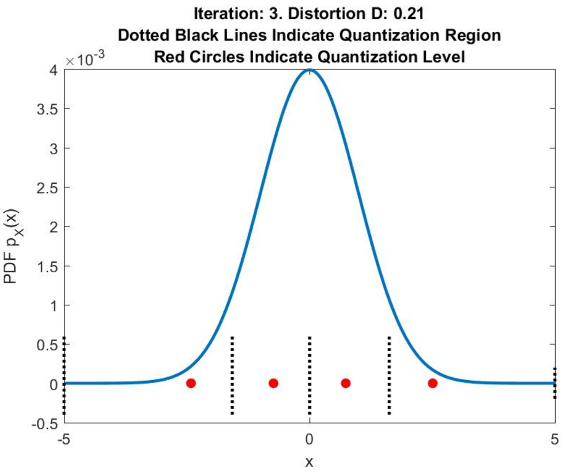
Lloyd-Max Algorithm

Gaussian PDF: Iteration 2



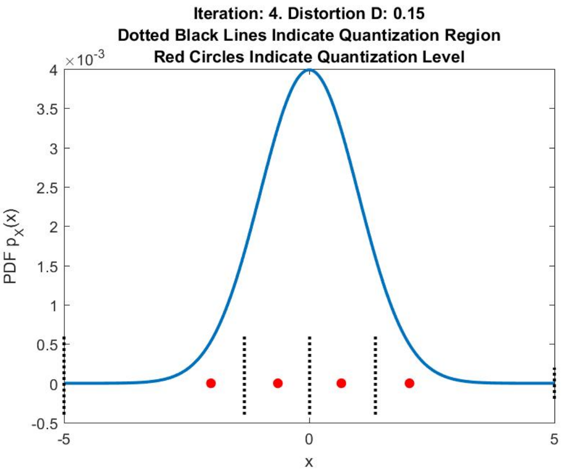
Lloyd-Max Algorithm

Gaussian PDF: Iteration 3



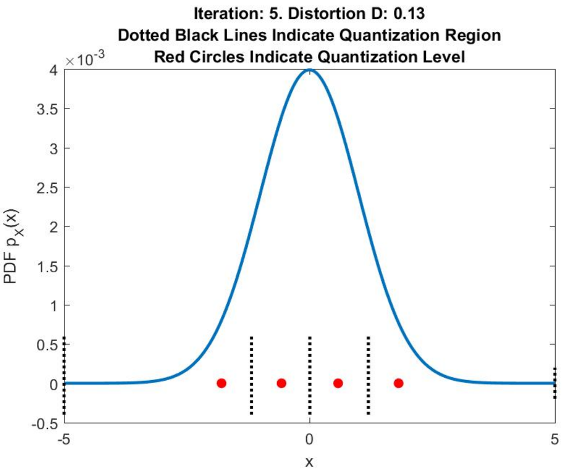
Lloyd-Max Algorithm

Gaussian PDF: Iteration 4



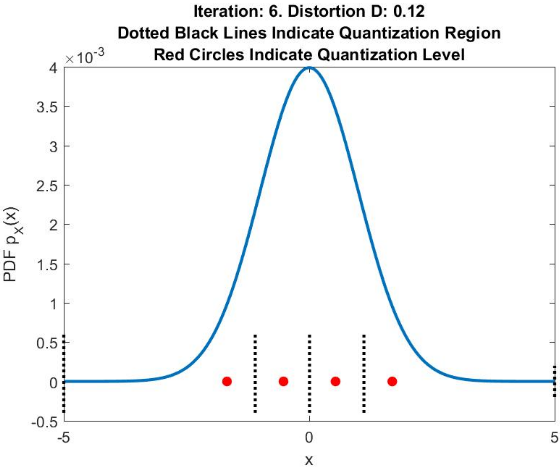
Lloyd-Max Algorithm

Gaussian PDF: Iteration 5



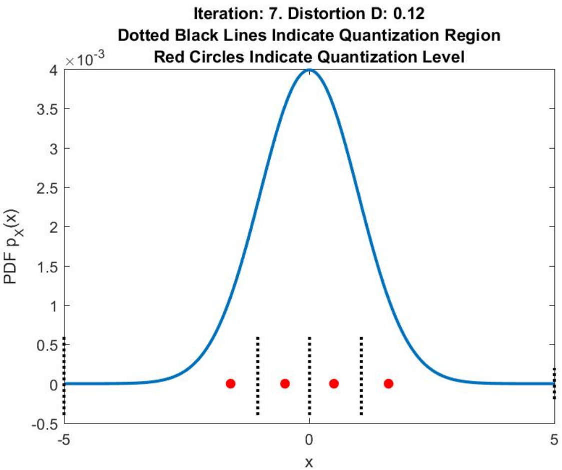
Lloyd-Max Algorithm

Gaussian PDF: Iteration 6



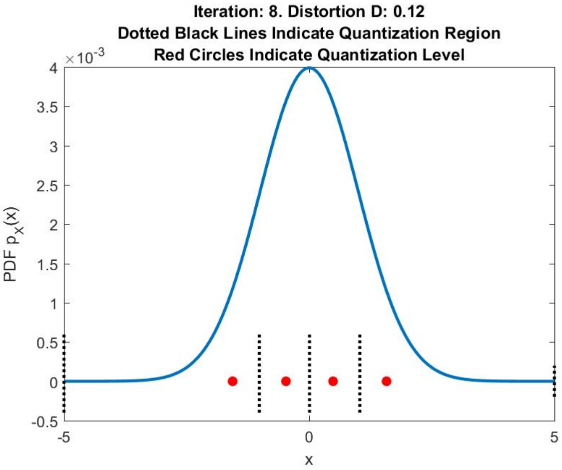
Lloyd-Max Algorithm

Gaussian PDF: Iteration 7



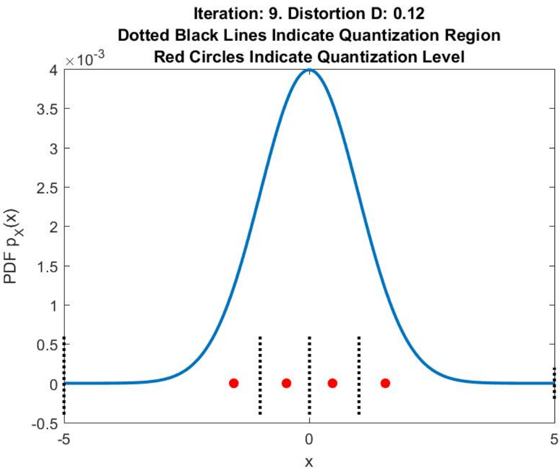
Lloyd-Max Algorithm

Gaussian PDF: Iteration 8



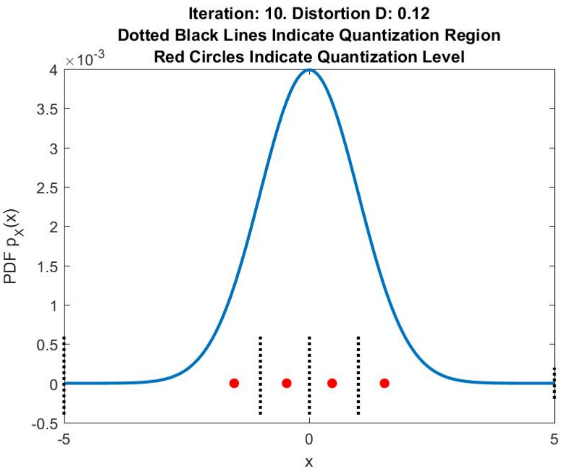
Lloyd-Max Algorithm

Gaussian PDF: Iteration 9



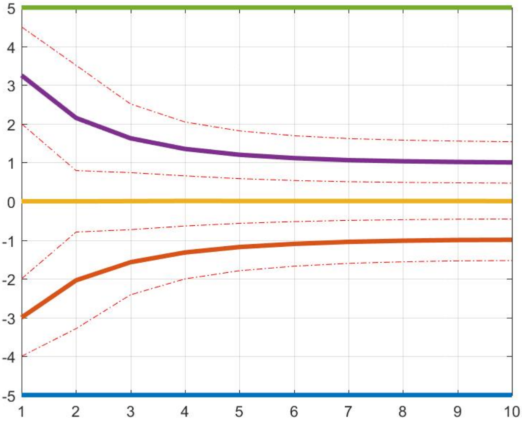
Lloyd-Max Algorithm

Gaussian PDF: Iteration 10



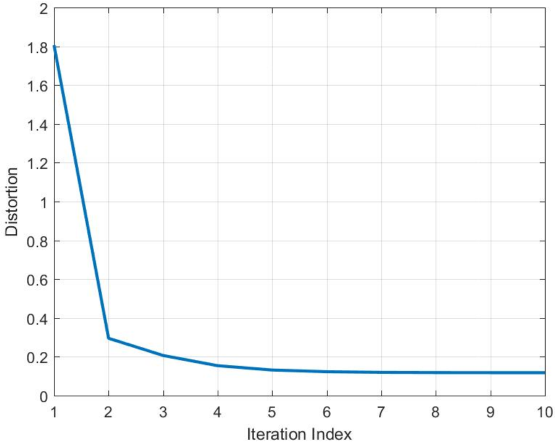
Lloyd-Max Algorithm

Gaussian PDF: Algorithm Convergence



Lloyd-Max Algorithm

Gaussian PDF: Algorithm Convergence



Lloyd-Max Algorithm
Gaussian PDF

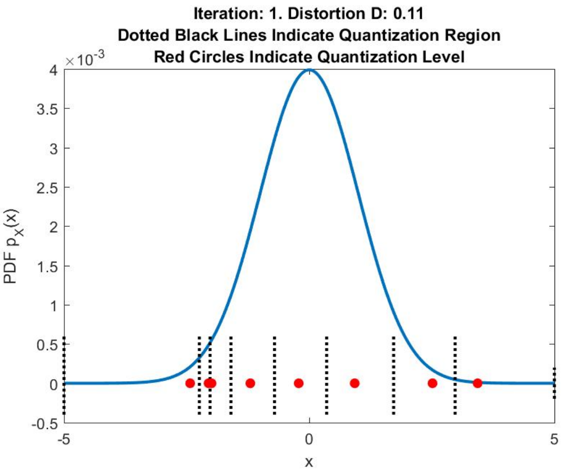
- Lloyd-Max algorithm is applied to the Gaussian PDF given as

$$p(x) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2}\right)$$

- $L = 8$, with initial quantization levels $\{\tilde{x}_1, \tilde{x}_2, \ldots, \tilde{x}_L\}$ set randomly.
- Initial quantization boundaries are found from $x_k = \frac{\tilde{x}_k + \tilde{x}_{k+1}}{2}, 1 \le k \le L-1.$
- Subsequent iterations of Lloyd-Max algorithm are depicted in the following slides.

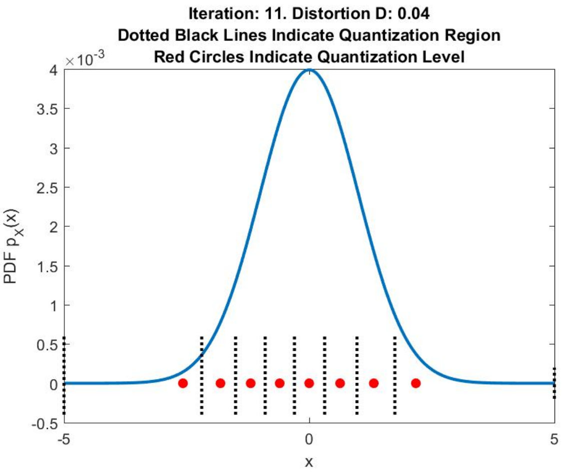
Lloyd-Max Algorithm

Gaussian PDF: Initialization (Iteration 1)



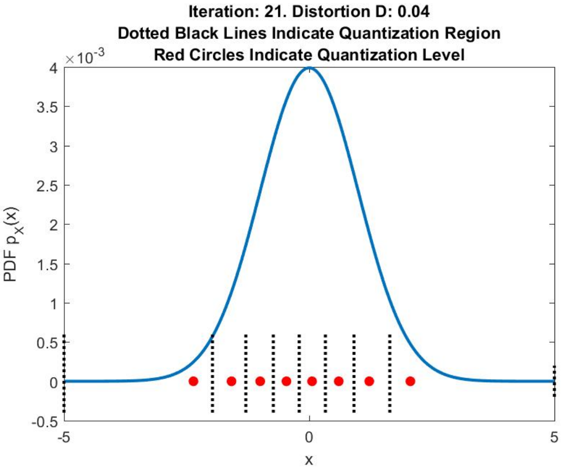
Lloyd-Max Algorithm

Gaussian PDF: Iteration 11



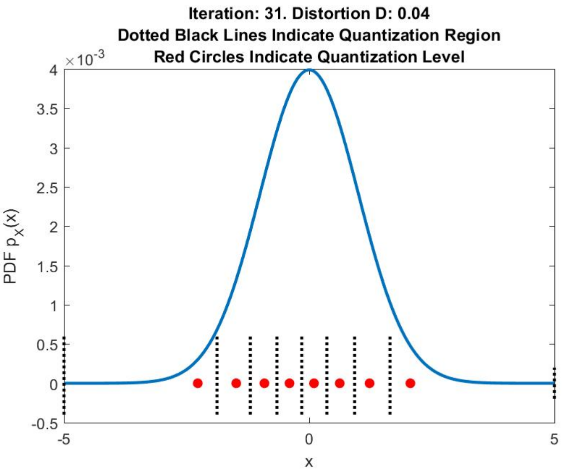
Lloyd-Max Algorithm

Gaussian PDF: Iteration 21



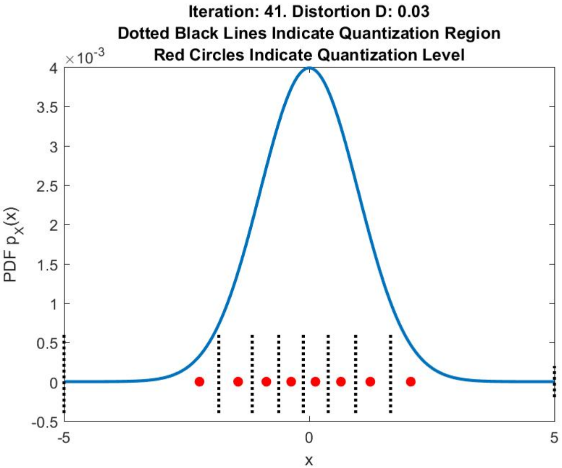
Lloyd-Max Algorithm

Gaussian PDF: Iteration 31



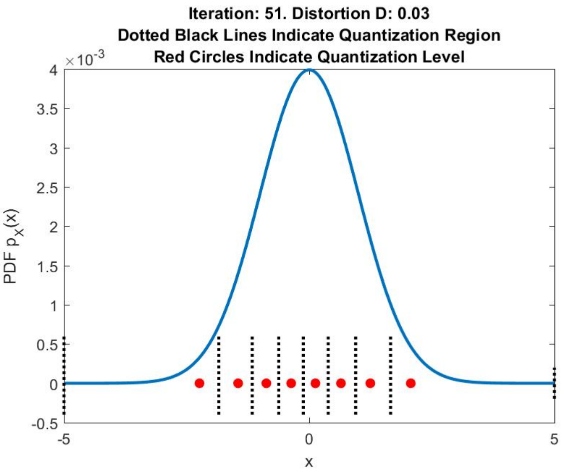
Lloyd-Max Algorithm

Gaussian PDF: Iteration 41



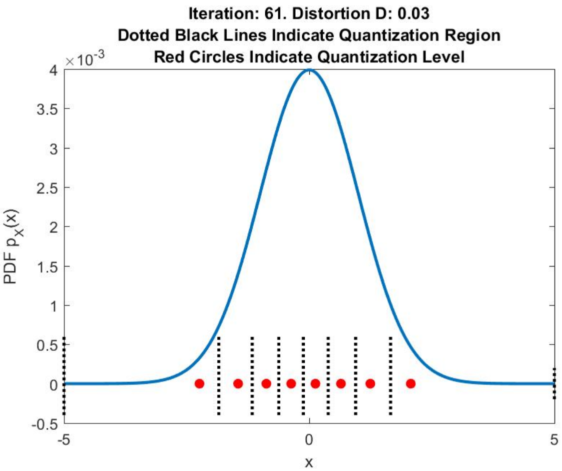
Lloyd-Max Algorithm

Gaussian PDF: Iteration 51



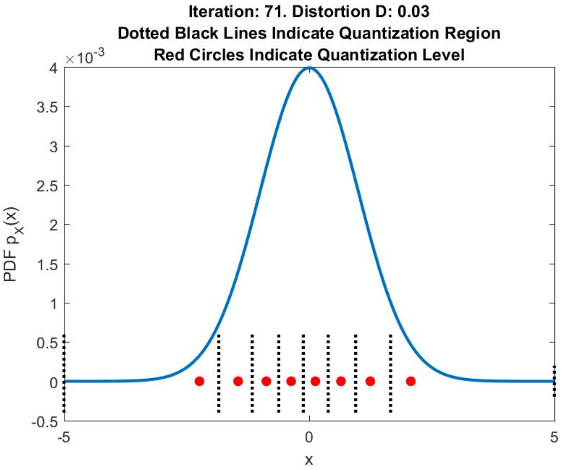
Lloyd-Max Algorithm

Gaussian PDF: Iteration 61



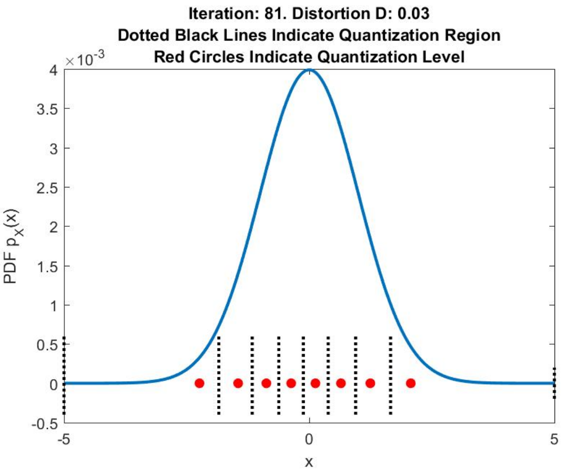
Lloyd-Max Algorithm

Gaussian PDF: Iteration 71



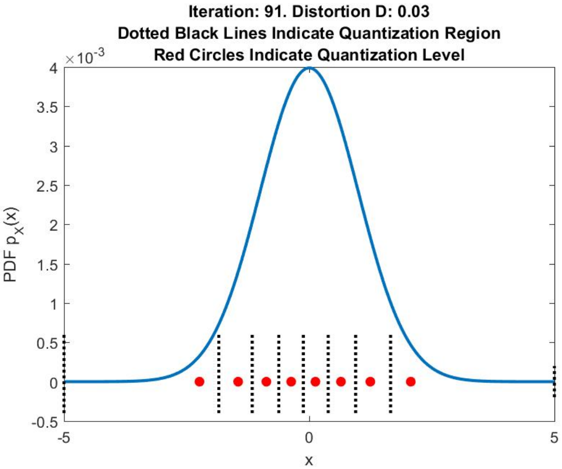
Lloyd-Max Algorithm

Gaussian PDF: Iteration 81



Lloyd-Max Algorithm

Gaussian PDF: Iteration 91



Notes

- General method for finding the optimum nonuniform quantizer
  - $\rightarrow$ A key requirement or limitation: requires the knowledge of PDF of the data.
- May take many iterations to converge depending on the initial seed
- It is possible that Lloyd-Max algorithm converges to local optimum instead of global optimum

Summary

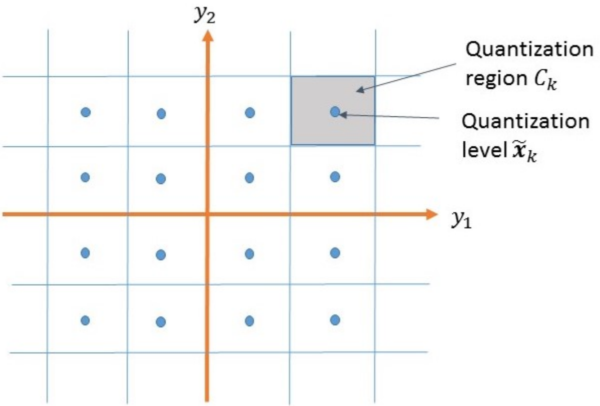
- Basic idea: concentrate quantization levels in areas of large PDF
  - → Most high speed analog to digital converters implement uniform quantization. Companding can be used to approximate nonuniform quantization.
- Non-uniform quantization can improve SNR by several dBs when the samples have a high dynamic range.
- However, scalar quantization still performs several dBs worse than Rate-Distortion bound
- Solution: Vector Quantization

Vector Quantization

- Quantize blocks of n samples at a time $\mathbf{Y} = \{y_1, y_2, \ldots, y_n\}$, where
  - $\rightarrow$ n is the *dimension* of the quantizer
  - $\rightarrow$ L quantization levels $\{\tilde{x}_k, k = 1, \ldots, L\}$ are replaced by L quantization vectors $\{\tilde{\mathbf{x}}_\mathbf{k}, k = 1, \ldots, L\}$
  - $\rightarrow$ Quantization intervals $((x_k, x_{k+1}), k = 1, \ldots, L)$ are replaced by quantization regions $(C_k, k = 1, \ldots, L), C_k \in R^n$
- Functional form of vector quantization: $\tilde{\mathbf{x}}_\mathbf{k} = f_Q(\mathbf{Y})$, if $\mathbf{Y} \in C_k$
- Rate of VQ $R = \dfrac{\log_2 L}{n}$ bits per sample (fractional rates are now possible)

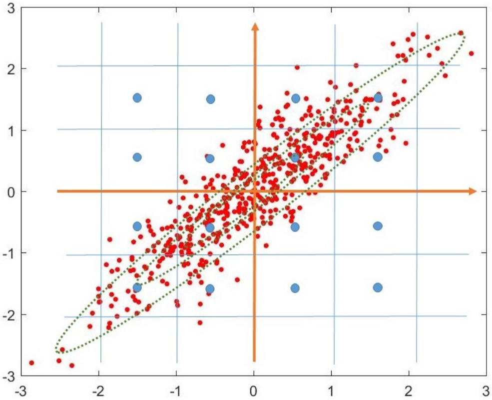
Example VQ

with $L = 16$ and $n = 2$



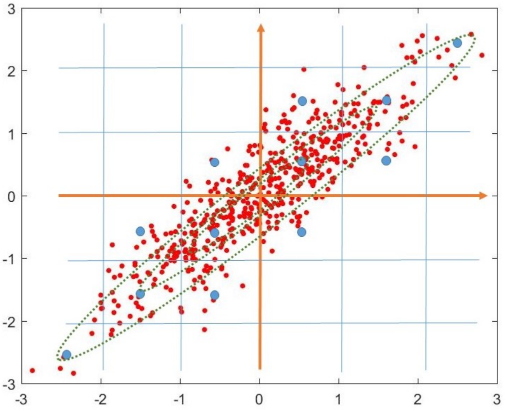
Why VQ Works?

- While scalar quantization can be wasteful of bits with correlated samples,. . .



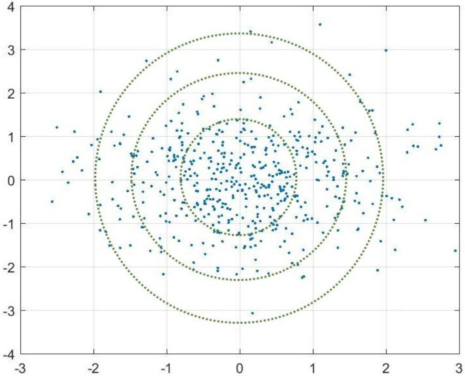
Why VQ Works?

- . . . VQ is better able to exploit correlation between samples



Why VQ Works?

- VQ gives better shaped quantization regions even in absence of correlation
- A shaping gain of about 1.5 dB is possible with VQ



Speech Encoding

- All speech coding techniques require quantization.
- Many of them employ additional properties of speech

[Temporal Waveform Coding:] attempts to represent time domain samples of speech

[Spectral Waveform Coding:] attempts to represent spectral characteristics of speech

[Model based Coding:] replicates a model of the process by which the Human auditory tract generates the speech

Speech Encoding

- All speech coding techniques require quantization.
- Many of them employ additional properties of speech

[Temporal Waveform Coding:] attempts to represent time domain samples of speech

[Spectral Waveform Coding:] attempts to represent spectral characteristics of speech

[Model based Coding:] replicates a model of the process by which the Human auditory tract generates the speech

Speech Encoding

- All speech coding techniques require quantization.
- Many of them employ additional properties of speech

  [Temporal Waveform Coding:] attempts to represent time domain samples of speech

  [Spectral Waveform Coding:] attempts to represent spectral characteristics of speech

  [Model based Coding:] replicates a model of the process by which the Human auditory tract generates the speech

Temporal Waveform Coding

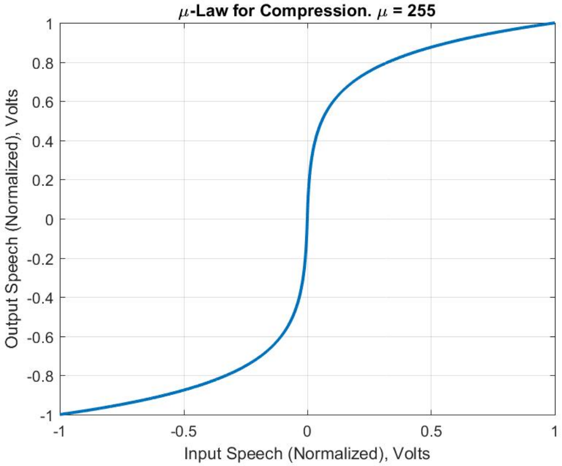
1. *Pulse Code Modulation (PCM):* directly quantizes the samples of speech
   - *Companding:*
     - In human speech, probability of low amplitudes is greater than that of the high amplitudes.
     - Need to use non-uniform quantization.
     - An alternate is nonlinear compander followed by the uniform quantizer
     - Nonlinearity for μ-law companding: $y = \dfrac{\log(1+\mu|x|)}{\log(1+\mu)}$
   - Example: 64 kbps PCM for telephone lines:
     - Speech signal (with a bandwidth of 4 kHz) is sampled at 8 ksps
     - Encodes each sample with 8 bit uniform quantizer
     - Uses companding with $\mu = 255$
2. *Differential PCM (DPCM):* speech samples are strongly correlated from one sampling instant to the next. DPCM quantizes the difference between the current sample and the predicted value of this sample using the past samples.

Temporal Waveform Coding

1. *Pulse Code Modulation (PCM):* directly quantizes the samples of speech
   - *Companding:*
     - → In human speech, probability of low amplitudes is greater than that of the high amplitudes.
     - → Need to use non-uniform quantization.
     - → An alternate is nonlinear compander followed by the uniform quantizer
     - → Nonlinearity for μ-law companding: $y = \dfrac{\log(1+\mu|x|)}{\log(1+\mu)}$
   - Example: 64 kbps PCM for telephone lines:
     - → Speech signal (with a bandwidth of 4 kHz) is sampled at 8 ksps
     - → Encodes each sample with 8 bit uniform quantizer
     - → Uses companding with $\mu = 255$
2. *Differential PCM (DPCM):* speech samples are strongly correlated from one sampling instant to the next. DPCM quantizes the difference between the current sample and the predicted value of this sample using the past samples.

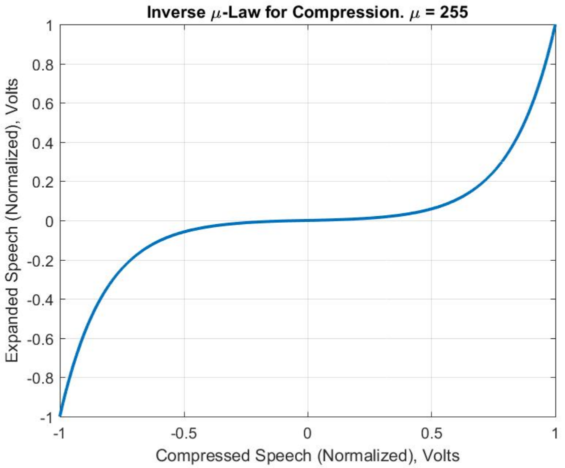
Speech Encoding

μ Law used at Speech Encoder



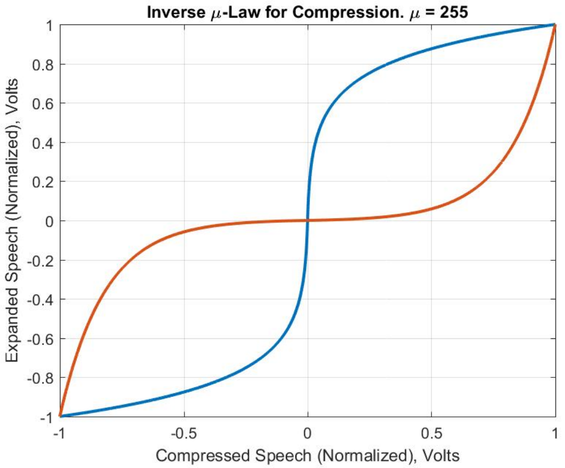
Speech Encoding

Inverse μ Law used at Speech Decoder



Speech Encoding

μ Law and Inverse μ Law



Temporal Waveform Coding

3. Adaptive PCDM
   - Perceptual quality is determined by the relative accuracy of the quantization to the size of speech
   - Adaptive quantizers vary the step size between the quantization levels depending on whether the speech is loud or soft
   - Adaptive DPCM at 32 kbps is a common and easily implemented, and used in Digital European Cordless Telephone or DECT standard
4. Delta Modulation
   - An extreme case of PCM in which signal is oversampled and $R = 1$ bit/sample, i.e., needs a very simple one bit quantizer
   - Equivalent to just knowing the zero-crossing of the speech signal
   - Adaptive Delta Modulation can produce reasonable quality speech at 16 kbps

Temporal Waveform Coding

3. Adaptive PCDM
   - Perceptual quality is determined by the relative accuracy of the quantization to the size of speech
   - Adaptive quantizers vary the step size between the quantization levels depending on whether the speech is loud or soft
   - Adaptive DPCM at 32 kbps is a common and easily implemented, and used in Digital European Cordless Telephone or DECT standard
4. Delta Modulation
   - An extreme case of PCM in which signal is oversampled and $R = 1$ bit/sample, i.e., needs a very simple one bit quantizer
   - Equivalent to just knowing the zero-crossing of the speech signal
   - Adaptive Delta Modulation can produce reasonable quality speech at 16 kbps

Spectral Waveform Coding

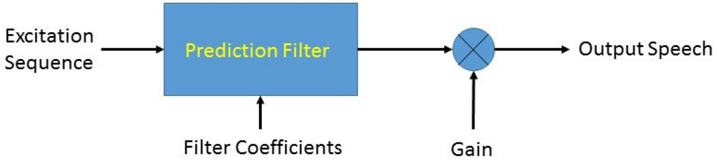
1. Subband Coding
   - Human perception of speech quality depends on the frequency band. Higher MSE may be tolerated at very low and very high frequencies
   - Subband coders filter the speech signal into multiple bands using Quadrature Mirror Filters (QMFs)
   - Filtered signals are quantied using standard PCM, with a different value of R for different bands
2. Adaptive Transform Coding
   - Correlated time domain signals are transformed into frequency domain samples using FFT or Discrete Cosine Transform
   - These frequency domain signals are represented using their perceptual importance
   - Often combined with VQ and Linear Prediction

Spectral Waveform Coding

1. Subband Coding
   - Human perception of speech quality depends on the frequency band. Higher MSE may be tolerated at very low and very high frequencies
   - Subband coders filter the speech signal into multiple bands using Quadrature Mirror Filters (QMFs)
   - Filtered signals are quantied using standard PCM, with a different value of R for different bands
2. Adaptive Transform Coding
   - Correlated time domain signals are transformed into frequency domain samples using FFT or Discrete Cosine Transform
   - These frequency domain signals are represented using their perceptual importance
   - Often combined with VQ and Linear Prediction

Linear Predictive Speech Coding

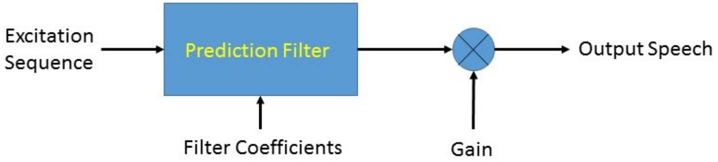
- Human speech is modeled as noise (air from the lungs) exciting a linear filter (throat, vocal cords and the mouth)
- Following are quantized and transmitted to the receiver: excitation sequence, filter coefficients, gain



- VSELP (Vector Sum Excited Linear Prediction)
  - → 20 ms frames with 159 bits per frame. Data rate of approx 8 kbps.
  - → Two stage VQ is used to quantize the excitation sequence
  - → Bits (such as those representing the filter gain) are more critical for perceptual quality. These are protected by error correction coding

Linear Predictive Speech Coding

- Human speech is modeled as noise (air from the lungs) exciting a linear filter (throat, vocal cords and the mouth)
- Following are quantized and transmitted to the receiver: excitation sequence, filter coefficients, gain



- VSELP (Vector Sum Excited Linear Prediction)
  - → 20 ms frames with 159 bits per frame. Data rate of approx 8 kbps.
  - → Two stage VQ is used to quantize the excitation sequence
  - → Bits (such as those representing the filter gain) are more critical for perceptual quality. These are protected by error correction coding

Relative Performance

Table: Vocoder Comparison

Type	Rate (kbps)	Complexity (MIPS)	Delay (ms)	Quality
PCM	64	0.01	0	High
ADPCM	32	0.1	0	High
Subband	16	1	25	High
VSELP	8	$\sim$100	35	Fair

- How to measure the quality or the distortion of the speech encoder and quantizers?
  - MSE is a possible measure, however, it does not relate to how good the encoded and quantized speech sounds to the human ear
  - Alternative distortion measures: Perceptually Weighted MSE, Segmental SNR, Itakura-Saito, Log Spectral Distance, etc.
  - No single measure has been found to accurately quantify the perceptual quality
  - Subjective assessment of speech quality involves Mean Opinion Score (MOS) testing: subjects rate speech on a scale of 1 (unintelligible) to 5 (perfect). Toll quality telephone speech rates around 4.3.

Relative Performance

Table: Vocoder Comparison

Type	Rate (kbps)	Complexity (MIPS)	Delay (ms)	Quality
PCM	64	0.01	0	High
ADPCM	32	0.1	0	High
Subband	16	1	25	High
VSELP	8	$\sim$100	35	Fair

- How to measure the quality or the distortion of the speech encoder and quantizers?
  - MSE is a possible measure, however, it does not relate to how good the encoded and quantized speech sounds to the human ear
  - Alternative distortion measures: Perceptually Weighted MSE, Segmental SNR, Itakura-Saito, Log Spectral Distance, etc.
  - No single measure has been found to accurately quantify the perceptual quality
  - Subjective assessment of speech quality involves Mean Opinion Score (MOS) testing: subjects rate speech on a scale of 1 (unintelligible) to 5 (perfect). Toll quality telephone speech rates around 4.3.

References

- N. S. Jayant, "Coding Speech at Low Bit Rates," IEEE Spectrum, August 1986.
- N. S. Jayant, et al., "Coding of Speech and Wideband Audio," AT&T Technical Journal, October 1990.