ECE 729

Vector Quantization, Voronoi Regions, Linear Estimation, and the Lloyd–Max Algorithm

1. Vector Quantizers

Let $\boldsymbol{X}$ be an $\mathbb{R}^d$-valued random vector. An ***N*-level vector quantizer** is any function q of the form

$$q(\boldsymbol{x}) := \sum_{i=1}^{N} \boldsymbol{c}_i I_{V_i}(\boldsymbol{x}), \tag{1}$$

where the sets $V_1, \ldots, V_N$ form a partition of $\mathbb{R}^d$ and the $\boldsymbol{c}_i \in \mathbb{R}^d$.

1.1. Optimizing the Partition – Voronoi Regions

If the $\boldsymbol{c}_i$ are given, what is the best choice of partition to minimize the mean-squared error $\mathsf{E}[\|\boldsymbol{X} - q(\boldsymbol{X})\|^2]$?

We claim that no partition can do better than the partition of **Voronoi regions** defined by[1]

$$V_j^* := \{\boldsymbol{x} \in \mathbb{R}^d : \|\boldsymbol{x} - \boldsymbol{c}_j\|^2 \le \|\boldsymbol{x} - \boldsymbol{c}_i\|^2 \text{ for all } i\}. \tag{2}$$

To establish this claim, let

$$q^*(\boldsymbol{x}) := \sum_{j=1}^{n} \boldsymbol{c}_j I_{V_j^*}(\boldsymbol{x}).$$

Then write

$$\begin{aligned}
\mathsf{E}[\|\boldsymbol{X} - q(\boldsymbol{X})\|^2] &= \mathsf{E}[\|\boldsymbol{X} - q(\boldsymbol{X})\|^2 \cdot 1 \cdot 1] \\
&= \mathsf{E}\left[\|\boldsymbol{X} - q(\boldsymbol{X})\|^2 \left(\sum_{i=1}^{N} I_{V_i}(\boldsymbol{X})\right)\left(\sum_{j=1}^{N} I_{V_j^*}(\boldsymbol{X})\right)\right] \\
&= \sum_{i=1}^{N}\sum_{j=1}^{N} \mathsf{E}[\|\boldsymbol{X} - q(\boldsymbol{X})\|^2 I_{V_i}(\boldsymbol{X}) I_{V_j^*}(\boldsymbol{X})] \\
&= \sum_{i=1}^{N}\sum_{j=1}^{N} \mathsf{E}[\|\boldsymbol{X} - \boldsymbol{c}_i\|^2 I_{V_i}(\boldsymbol{X}) I_{V_j^*}(\boldsymbol{X})],
\end{aligned}$$

since $q(\boldsymbol{x}) := \boldsymbol{c}_i$ for $\boldsymbol{x} \in V_i$. On the other hand, for $\boldsymbol{x} \in V_j^*$, we know that

$$\|\boldsymbol{x} - \boldsymbol{c}_j\|^2 \le \|\boldsymbol{x} - \boldsymbol{c}_i\|^2, \quad \text{for all } i.$$

Hence,

$$\begin{aligned}
\mathsf{E}[\|\boldsymbol{X} - q(\boldsymbol{X})\|^2] &\ge \sum_{i=1}^{N}\sum_{j=1}^{N} \mathsf{E}[\|\boldsymbol{X} - \boldsymbol{c}_j\|^2 I_{V_i}(\boldsymbol{X}) I_{V_j^*}(\boldsymbol{X})] \\
&= \sum_{i=1}^{N}\sum_{j=1}^{N} \mathsf{E}[\|\boldsymbol{X} - q^*(\boldsymbol{X})\|^2 I_{V_i}(\boldsymbol{X}) I_{V_j^*}(\boldsymbol{X})] \\
&= \mathsf{E}\left[\|\boldsymbol{X} - q^*(\boldsymbol{X})\|^2 \left(\sum_{i=1}^{N} I_{V_i}(\boldsymbol{X})\right)\left(\sum_{j=1}^{N} I_{V_j^*}(\boldsymbol{X})\right)\right] \\
&= \mathsf{E}[\|\boldsymbol{X} - q^*(\boldsymbol{X})\|^2].
\end{aligned}$$

[1] To be precise, we should let B_j denote the set in (2) and then put $V_1^* := B_1$ and $V_j^* := B_j \cap B_{j-1}^c \cap \cdots \cap B_1^c$ for $j = 2, \ldots, N$.

1.2. Optimizing the c_i – Linear Estimation

If the sets V_i of an arbitrary quantizer q in (1) are given, what are the best $\boldsymbol{c}_i$ to use?

Let $\boldsymbol{Y} := [I_{V_1}(\boldsymbol{X}), \ldots, I_{V_N}(\boldsymbol{X})]'$. Then $q(\boldsymbol{X}) = A\boldsymbol{Y}$, where A is the $d \times N$ matrix $A := [\boldsymbol{c}_1, \ldots, \boldsymbol{c}_N]$. Hence, the mean-squared error is $\mathsf{E}[\|X - A\boldsymbol{Y}\|^2]$. In other words, finding the best vectors $\boldsymbol{c}_i$ is equivalent to finding the best linear transformation (matrix) A. The best matrix is any solution of the **normal equations**

$$AR_{\boldsymbol{Y}} = R_{\boldsymbol{XY}}, \tag{3}$$

where $R_{\boldsymbol{Y}}$ is the correlation (not covariance) matrix $R_{\boldsymbol{Y}} := \mathsf{E}[\boldsymbol{YY}']$ and $R_{\boldsymbol{XY}}$ is the cross-correlation matrix $R_{\boldsymbol{XY}} := \mathsf{E}[\boldsymbol{XY}']$.

2. The Lloyd–Max Algorithm

The repeated alternating iteration of optimizing the partition and $\boldsymbol{c}_i$ is the **Lloyd–Max algorithm** [1, 2]. The scalar case, $d = 1$, is quite simple. First, if $c_1 < \cdots < c_N$, then (2) reduces to

$$V_j^* = \left(\frac{c_{j-1} + c_j}{2}, \frac{c_j + c_{j+1}}{2}\right].$$

Second, since R_Y is diagonal, (3) is easily solved to yield

$$c_j^* = \frac{\mathsf{E}[X I_{V_j}(X)]}{\mathsf{P}(X \in V_j)}.$$

References

[1] S. P. Lloyd, "Least squares quantization in PCM," unpublished Bell Laboratories Memorandum, July 31, 1957; also *IEEE Trans. Inform. Theory*, vol. IT-28, pp. 129–137, Mar. 1982.

[2] J. Max, "Quantizing for minimum distortion," *IRE Trans. Inform. Theory*, vol. IT-6, pp. 7–12, Mar. 1960.

[3] A. V. Trushkin, "Sufficient conditions for uniqueness of a locally optimal quantizer for a class of convex error weighting functions," *IEEE Trans. Inform. Theory*, vol. IT-28, no. 2, pp. 187–198, Mar. 1982.