# VERTICAL (CAC) STAGE-CODEVECTORS

BARNES 2007

$$e = \sum_{k=1}^{K} \sum_{\substack{i=1 \\ x_i \in \mathcal{K}_k}}^{N_{\mathcal{K}_k}} (x_i - \hat{m}_k)^2$$

expand summation for easier understanding

$$= \sum_{\substack{i=1 \\ x_i \in \mathcal{K}_1}}^{N_{\mathcal{K}_1}} (x_i - \hat{m}_1)^2 p_X(x_i) + \sum_{\substack{i=1 \\ x_i \in \mathcal{K}_2}}^{N_{\mathcal{K}_2}} (x_i - \hat{m}_2)^2 p_X(x_i) + \ldots + \sum_{\substack{i=1 \\ x_i \in \mathcal{K}_K}}^{N_{\mathcal{K}_K}} (x_i - \hat{m}_K)^2 p_X(x_i)$$

here, there is a derivative of every equivalent codevector w.r.t. $z_{\alpha,\beta}$

$$\frac{\partial e}{\partial z_{\alpha,\beta}} = \sum_{\substack{i=1 \\ x_i \in \mathcal{K}_1}}^{N_{\mathcal{K}_1}} 2(x_i - \hat{m}_1) \frac{\partial \hat{m}_1}{\partial z_{\alpha,\beta}} p_X(x_i) + \sum_{\substack{i=1 \\ x_i \in \mathcal{K}_2}}^{N_{\mathcal{K}_2}} 2(x_i - \hat{m}_2) \frac{\partial \hat{m}_2}{\partial z_{\alpha,\beta}} p_X(x_i) + \ldots + \sum_{\substack{i=1 \\ x_i \in \mathcal{K}_K}}^{N_{\mathcal{K}_K}} 2(x_i - \hat{m}_K) \frac{\partial \hat{m}_K}{\partial z_{\alpha,\beta}} p_X(x_i)$$

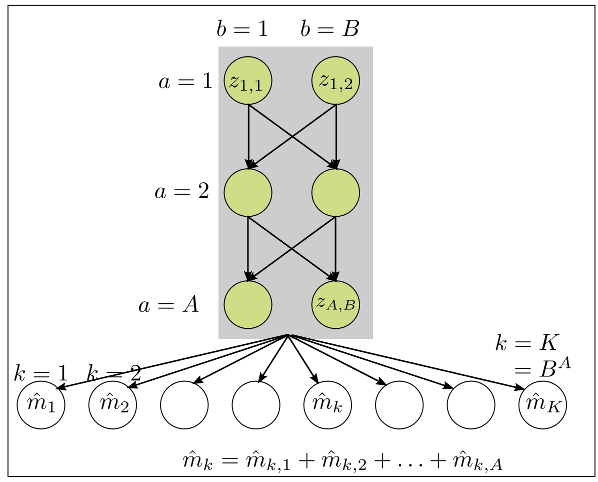
we consider only those data points $x_i$ that map to $z_{\alpha,\beta}$, i.e. $x_i \in \mathcal{H}_{\alpha,\beta}$, i.e. $\hat{m}_k = \sum_{\substack{a=1 \\ a \neq \alpha}}^{A} \hat{m}_{k,a} + z_{\alpha,\beta}$

$$= \sum_{k=1}^{K} \sum_{\substack{i=1 \\ x_i \in (\mathcal{H}_{\alpha,\beta} \cap \mathcal{K}_k)}}^{N_{\mathcal{H}_{\alpha,\beta} \cup \mathcal{K}_k}} 2(x_i - \hat{m}_k) \frac{\partial \hat{m}_k}{\partial z_{\alpha,\beta}} p_X(x_i)$$

$$= \sum_{k=1}^{K} \sum_{\substack{i=1 \\ x_i \in (\mathcal{H}_{\alpha,\beta} \cap \mathcal{K}_k)}}^{N_{\mathcal{H}_{\alpha,\beta} \cap \mathcal{K}_k}} 2 \left( x_i - \left[ \sum_{\substack{a=1 \\ a \neq \alpha}}^{A} \hat{m}_{k,a} + z_{\alpha,\beta} \right] \right) \underbrace{\frac{\partial \left[ \sum_{\substack{a=1 \\ a \neq \alpha}}^{A} \hat{m}_{k,a} + z_{\alpha,\beta} \right]}{\partial z_{\alpha,\beta}}}_{1} p_X(x_i)$$

let $\frac{\partial e}{\partial z_{\alpha,\beta}} = 0$

$$\Rightarrow z_{\alpha,\beta} = \frac{\sum_{k=1}^{K} \sum_{\substack{i=1 \\ x_i \in (\mathcal{H}_{\alpha,\beta} \cap \mathcal{K}_k)}}^{N_{\mathcal{H}_{\alpha,\beta} \cap \mathcal{K}_k}} \left( x_i - \left[ \sum_{\substack{a=1 \\ a \neq \alpha}}^{A} \hat{m}_{k,a} + z_{\alpha,\beta} \right] \right) p_X(x_i)}{\sum_{k=1}^{K} \sum_{\substack{i=1 \\ x_i \in (\mathcal{H}_{\alpha,\beta} \cap \mathcal{K}_k)}}^{N_{\mathcal{H}_{\alpha,\beta} \cap \mathcal{K}_k}} p_X(x_i)}$$




| | |
|---|---|
| $A$ | : number of stages |
| $B$ | : number of templates per stage |
| $K$ | : number of clusters= number of equivalent code-vectors<br>=number of paths through trellis=number of SoC descriptors = $B^A$ |
| $a$ | $= \{1, 2, \ldots, A\}$ |
| $b$ | $= \{1, 2, \ldots, B\}$ |
| $k$ | $= \{1, 2, \ldots, K\}$ |
| $e$ | : mean squared error between inputs and reconstructed inputs |
| $\mathcal{K}_k$ | : $k$-th cluster |
| $\hat{m}_k$ | : $k$-th mean, i.e. $k$-th equivalent code-vector (uses $k$-th SoC descriptor) |
| $\hat{m}_{k,a}$ | : stage code-vector, $a$-th additive component of $\hat{m}_k$ |
| $N_{\mathcal{K}_k}$ | : number of input data points that map to the $k$-th cluster |
| $z_{\alpha,\beta}$ | : stage-codevector at stage $a = \alpha$ and template $b = \beta$ |
| $\mathcal{H}_{\alpha,\beta}$ | : cluster of points that map to $z_{\alpha,\beta}$ stage-codevector |
| $N_{\mathcal{H}_{\alpha,\beta} \cup \mathcal{K}_k}$ | : number of data points from cluster $\mathcal{K}_k$ that also lie in cluster $\mathcal{H}_{\alpha,\beta}$ |

# HORIZONTAL-VERTICAL STAGE-CODEVECTORS

$$e = \sum_{k=1}^{K} \sum_{\substack{i=1 \\ x_i \in \mathcal{K}_k}}^{N_{\mathcal{K}_k}} (x_i - \hat{m}_k)^2$$

$$= \sum_{\substack{i=1 \\ x_i \in \mathcal{K}_1}}^{N_{\mathcal{K}_1}} (x_i - \hat{m}_1)^2 p_X(x_i) + \sum_{\substack{i=1 \\ x_i \in \mathcal{K}_2}}^{N_{\mathcal{K}_2}} (x_i - \hat{m}_2)^2 p_X(x_i) + \ldots + \sum_{\substack{i=1 \\ x_i \in \mathcal{K}_K}}^{N_{\mathcal{K}_K}} (x_i - \hat{m}_K)^2 p_X(x_i)$$

this step is the same as above

$$= \sum_{\substack{i=1 \\ x_i \in \mathcal{K}_1}}^{N_{\mathcal{K}_1}} \left( x_i - \sum_{a=1}^{A} \hat{m}_{1,a} \right)^2 p_X(x_i) + \sum_{\substack{i=1 \\ x_i \in \mathcal{K}_2}}^{N_{\mathcal{K}_2}} \left( x_i - \sum_{a=1}^{A} \hat{m}_{2,a} \right)^2 p_X(x_i) + \ldots + \sum_{\substack{i=1 \\ x_i \in \mathcal{K}_K}}^{N_{\mathcal{K}_K}} \left( x_i - \sum_{a=1}^{A} \hat{m}_{K,a} \right)^2 p_X(x_i)$$

we expand each equivalent codevector
in terms of its component stage-codevectors

- as a result of optimizing only one stage-codevector at a time, we optimize in causal anti-causal (vertical) direction only
- there is no horizontal optimization
- i.e., for a given stage, even though all stage-codevectors are held constant, there is no joint optimization between them

in CAC, of these stage-codevectors, we pick only those that are equal to $z_{\alpha,\beta}$

$$= \sum_{\substack{i=1 \\ x_i \in \mathcal{K}_1}}^{N_{\mathcal{K}_1}} \left( x_i - \sum_{\substack{a=1 \\ a \neq \alpha}}^{A} \hat{m}_{1,a} - \hat{m}_{1,\alpha} \right)^2 p_X(x_i) + \sum_{\substack{i=1 \\ x_i \in \mathcal{K}_2}}^{N_{\mathcal{K}_2}} \left( x_i - \sum_{\substack{a=1 \\ a \neq \alpha}}^{A} \hat{m}_{2,a} - \hat{m}_{2,\alpha} \right)^2 p_X(x_i) + \ldots + \sum_{\substack{i=1 \\ x_i \in \mathcal{K}_K}}^{N_{\mathcal{K}_K}} \left( x_i - \sum_{\substack{a=1 \\ a \neq \alpha}}^{A} \hat{m}_{K,a} - \hat{m}_{K,\alpha} \right)^2 p_X(x_i)$$

- if instead of picking one stage-codevector at a time as in CAC, we pick all same-stage-codevectors, we get a formulation in which we can jointly optimize horizontally along a stage, and vertically as in CAC
- in CAC, optimization is done using a derivative wrt to a single stage-codevector
- the vertical optimization is done separately for every same-stage-codevector
- optimization of one stage-codevector is completely independent of optimizations for other same-stage-codevectors
- with the formulation below, we use GLA on vertical (causal anti-causal) residuals, to optimize all same-stage-codevectors simultaneously
- this is expected to give better results

$$= \sum_{\substack{i=1 \\ x_i \in \mathcal{K}_1}}^{N_{\mathcal{K}_1}} (g_i - \hat{m}_{1,\alpha})^2 p_X(x_i) + \sum_{\substack{i=1 \\ x_i \in \mathcal{K}_2}}^{N_{\mathcal{K}_2}} (g_i - \hat{m}_{2,\alpha})^2 p_X(x_i) + \ldots + \sum_{\substack{i=1 \\ x_i \in \mathcal{K}_K}}^{N_{\mathcal{K}_K}} (g_i - \hat{m}_{K,\alpha})^2 p_X(x_i)$$

here we are optimizing all stage-codevectors at stage $\alpha$