4. Machine Learning and Binaural Hearing

1. Binaural Features
2. Mapping Sounds onto Their Directions
3. Collecting Training Data
4. The Binaural Manifold
5. Localization with a Look-up Table
6. Linear Regression
7. **Piecewise Linear Regression**
8. Predicting the Direction of Speech
9. Principles of Sound Separation
10. Separation & Localization Method

Regression

- Reversed **Linear regression**, $\boldsymbol{x} \in \mathbb{R}^2$, $\boldsymbol{y} \in \mathbb{R}^D$:

$$\boldsymbol{y} = \mathbf{A}\boldsymbol{x} + \boldsymbol{b}, \quad \mathbf{A} \in \mathbb{R}^{D\times 2}, \boldsymbol{b} \in \mathbb{R}^D$$

- **Piecewise linear regression** (there are K possible mappings):

$$\boldsymbol{y} = \sum_{k=1}^{K} \mathbb{I}(z = k)(\mathbf{A}_k \boldsymbol{x} + \boldsymbol{b}_k + \boldsymbol{e}_k)$$

- $\mathbb{I}(z)$ is called an **indicator function**, that selects the k-th affine transformation $\mathbf{A}_k, \boldsymbol{b}_k$.
- $\boldsymbol{e}_k \in \mathbb{R}^D$ is an error vector accounting for the piecewise linear approximation.

Probabilistic Setting

- The joint input-output probability is decomposed:

$$p(\boldsymbol{y}, \boldsymbol{x}) = \sum_{k=1}^{K} p(\boldsymbol{y}|\boldsymbol{x}, Z = k)p(\boldsymbol{x}|Z = k)p(Z = k)$$

- Assuming Gaussian (normal) distributions we have:

$$\begin{aligned} p(\boldsymbol{y}|\boldsymbol{x}, Z = k) &= \mathcal{N}(\boldsymbol{y}; \mathbf{A}_k \boldsymbol{x} + \boldsymbol{b}_k, \boldsymbol{\Sigma}) \\ p(\boldsymbol{x}|Z = k) &= \mathcal{N}(\boldsymbol{x}; \boldsymbol{c}_k, \boldsymbol{\Gamma}_k) \\ p(Z = k) &= \pi_k \end{aligned}$$

Gaussian Mixture Model

- This formulation belongs to Gaussian Mixture Models (GMM).
- The model parameters are:

$$\theta = \{\boldsymbol{c}_k, \boldsymbol{\Gamma}_k, \pi_k, \mathbf{A}_k, \boldsymbol{b}_k, \boldsymbol{\Sigma}_k\}_{k=1}^{K}$$

- The parameters can be estimated via an expectation-maximization (EM) algorithm:
  1. Initialize the model parameters $\theta^{(0)} = \{\boldsymbol{c}_k^{(0)}, \boldsymbol{\Gamma}_k^{(0)}, \pi_k^{(0)}, \mathbf{A}_k^{(0)}, \boldsymbol{b}_k^{(0)}, \boldsymbol{\Sigma}_k^{(0)}\}_{k=1}^{K}$,
  2. Evaluate the posterior probabilities $r_{kn}^{(i)} = p(Z_n = k | \boldsymbol{x}_n, \boldsymbol{y}_n; \theta^{(i-1)})$,
  3. Maximize the **complete-data expected log-likelihood**:

$$\theta^{(i)} = \underset{\theta}{\text{argmax}}\, \mathrm{E}\big[\log p(\boldsymbol{X}, \boldsymbol{Y}, Z | \theta; \theta^{(i-1)})\big].$$

  4. Iterate steps 2 & 3 until convergence.

Posterior Probabilities (I)

- The optimal parameters $\hat{\theta}$ thus obtained allow to estimate the probability of $\boldsymbol{y}$ given $\boldsymbol{x}$:

$$p(\boldsymbol{y}|\boldsymbol{x};\hat{\theta}) = \sum_{k=1}^{K} \frac{\pi_k \mathcal{N}(\boldsymbol{x};\hat{\boldsymbol{c}}_k,\hat{\boldsymbol{\Gamma}}_k)}{\sum_{j=1}^{K} \pi_j \mathcal{N}(\boldsymbol{x};\hat{\boldsymbol{c}}_j,\hat{\boldsymbol{\Gamma}}_j)} \mathcal{N}(\boldsymbol{y};\hat{\mathbf{A}}_k\boldsymbol{x}+\hat{\boldsymbol{b}}_k,\hat{\boldsymbol{\Sigma}}_k).$$

Posterior Probabilities (II)

- More interesting, one can also evaluate the posterior probability of the sound direction, $\boldsymbol{x}$, given an ILPD vector, $\boldsymbol{y}$:

$$p(\boldsymbol{x}|\boldsymbol{y};\tilde{\theta}) = \sum_{k=1}^{K} \underbrace{\frac{\pi_k \mathcal{N}(\boldsymbol{y};\tilde{\boldsymbol{c}}_k,\tilde{\boldsymbol{\Gamma}}_k)}{\sum_{j=1}^{K}\pi_j \mathcal{N}(\boldsymbol{y};\tilde{\boldsymbol{c}}_j,\tilde{\boldsymbol{\Gamma}}_j)}}_{\tilde{\nu}_k(\text{proportion})} \mathcal{N}(\boldsymbol{x};\underbrace{\tilde{\boldsymbol{A}}_k\boldsymbol{y}+\tilde{\boldsymbol{b}}_k}_{\tilde{\boldsymbol{\mu}}_k(\text{mean})},\tilde{\boldsymbol{\Sigma}}_k)$$

$$= \sum_{k=1}^{K} \tilde{\nu}_k \mathcal{N}(\boldsymbol{x};\tilde{\boldsymbol{\mu}}_k,\tilde{\boldsymbol{\Sigma}}_k)$$

- The parameters $\tilde{\theta}$ are closed-form expressions of $\hat{\theta}$. [1]

[1] https://hal.inria.fr/hal-00863468/en

Session Summary

- Probabilistic treatment of regression
- Gaussian mixture model for regression
- Parameter estimation
- Bayes inversion