Introduction to Mixture Models

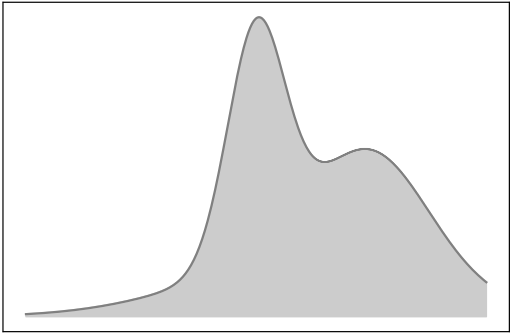
Yassine Yousfi

EECE 566

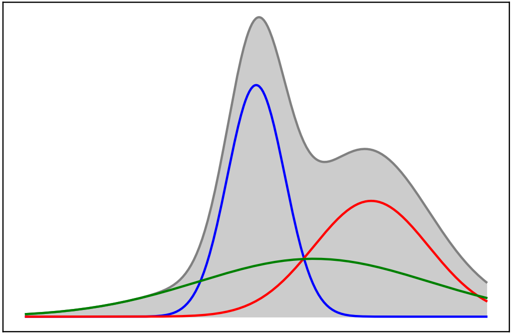
Introduction

- Real life measurements are usually more complex than simple distributions we studied (Gaussian, Laplace, etc.)
  - Can be multimodal
  - Can come from multiple sensors with different precisions
- We will extend to more complex probability distributions: **Mixture models**

Finite and countable mixture models



Finite and countable mixture models



Finite and countable mixture models

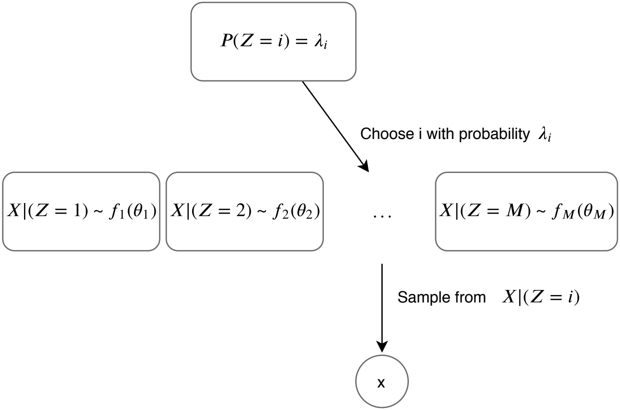
- Let $\lambda_1, \ldots, \lambda_M \in [0,1]$, verifying $\sum_i \lambda_i = 1$. And $f_i(\theta_i)$ potentially different types of probability distributions parametrized by θ_i. A mixture model can be written as:

$$p(x; \theta_1, \ldots, \theta_M, \lambda_1, \ldots, \lambda_M) = \sum_{i=1}^{M} \lambda_i f_i(x; \theta_i)$$

- Alternatively:

$$\begin{aligned} Z \in \{1, \ldots, M\} &\sim P(Z = i) = \lambda_i \\ X|Z &\sim f_{X|Z}(\theta_Z) \\ X &\sim \sum_z f_{X|Z}(X|Z = z; \theta_z) p(Z = z) = \sum_{i=1}^{M} \lambda_i f_i(\theta_i) \end{aligned}$$

Finite and countable mixture models



Finite and countable mixture models

- Is $f(x; \theta_1, \ldots, \theta_M, \lambda_1, \ldots, \lambda_M)$ a pdf? By definition ...

$$
\begin{aligned}
f(x; \theta_1, \ldots, \theta_M, \lambda_1, \ldots, \lambda_M) = \sum_{i=1}^{M} \lambda_i f_i(x; \theta_i) &\geq 0 \\
\int_{\mathcal{X}} f(x; \theta_1, \ldots, \theta_M, \lambda_1, \ldots, \lambda_M) dx &= \int_{\mathcal{X}} \sum_{i=1}^{M} \lambda_i f_i(x; \theta_i) \mathrm{d}x \\
&= \sum_{i=1}^{M} \lambda_i \underbrace{\int_{\mathcal{X}} f_i(x; \theta_i) dx}_{=1} \\
&= \sum_{i=1}^{M} \lambda_i = 1
\end{aligned}
$$

Example: finite Gaussian mixtures

- $p_i(\theta_i) = \mathcal{N}(\mu_i, \sigma_i^2)$:

$$\begin{aligned} f(x; \theta_1, \ldots, \theta_M, \lambda_1, \ldots, \lambda_M) &= \sum_{i=1}^{M} \lambda_i \mathcal{N}(x; \mu_i, \sigma_i^2) \\ &= \sum_{i=1}^{M} \lambda_i \frac{1}{\sqrt{2\pi\sigma_i^2}} \exp\left(\frac{-(x-\mu_i)^2}{2\sigma_i^2} \right) \end{aligned}$$

Uncountable mixture models

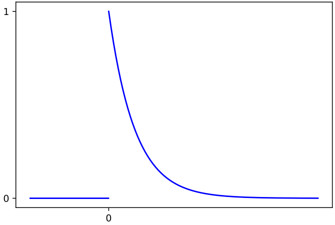
- We sample from a **continuous** set of parameters
- Prior knowledge $f(\lambda)$ of how these parameters have to be sampled (potentially parameterized)
- Simplifying assumption: same family of probability distribution model, parametrized from random samples of $f(\lambda)$:

$$\begin{aligned} Z &\sim f_Z(\lambda) \text{ prior} \\ X|Z &\sim f_{X|Z}(\theta_Z) \\ X &\sim \int_{\mathcal{Z}} f_{X|Z}(X|Z = z; \theta_z) f_Z(z; \lambda) \mathrm{d}z \end{aligned}$$

Example: Gaussian mixture with exponentially distributed variance

- We sample a variance σ^2 from an Exponential distribution:

$$f_{\sigma^2} = \begin{cases} e^{-t} & t \geq 0 \\ 0 & t < 0 \end{cases}$$



- And use this variance in $\mathcal{N}(0, \sigma^2)$

Example: Gaussian mixture with exponentially distributed variance

- $Z = \sigma^2$
- $f_Z(\lambda)$: Prior is exponential
- No λ: Prior is not parametrized
- $f_{X|Z}(\theta_Z) = \mathcal{N}(0, \sigma^2)$: Mixture elements are Gaussian r.v.'s

Example: Gaussian mixture with exponentially distributed variance

$$
\begin{aligned}
f(x|\sigma^2 = t) &= \frac{1}{\sqrt{2\pi t}} \exp\left(\frac{-x^2}{2t}\right) \\
f(x) &= \int_0^{+\infty} f(x|\sigma^2 = t) f_{\sigma^2}(t) \mathrm{d}t \\
&= \int_0^{+\infty} \frac{1}{\sqrt{2\pi t}} \exp\left(\frac{-x^2}{2t}\right) \exp(-t) \mathrm{d}t \\
&= \frac{1}{\sqrt{2\pi}} \int_0^{+\infty} \frac{1}{\sqrt{t}} \exp\left(\frac{-x^2}{2t} - t\right) \mathrm{d}t
\end{aligned}
$$

Example: Gaussian mixture with exponentially distributed variance

$$
\begin{aligned}
f(x) &= \frac{1}{\sqrt{2\pi}} \int_0^{+\infty} \frac{1}{\sqrt{t}} \exp\left(\frac{-x^2}{2t} - t\right) \mathrm{d}t \\
\text{(substitution) } y = \sqrt{t} : &= \frac{1}{\sqrt{2\pi}} \int_0^{+\infty} 2 \exp\left(\frac{-x^2}{2y^2} - y^2\right) \mathrm{d}y \\
\frac{x^2}{2y^2} + y^2 &= \left(y + \frac{|x|}{\sqrt{2}y}\right)^2 - \sqrt{2}x \\
z = y + \frac{x}{\sqrt{2}y} &: \quad \ldots
\end{aligned}
$$

Example: Gaussian mixture with exponentially distributed variance

$$
\begin{aligned}
f(x) &= \frac{1}{\sqrt{2\pi}} \int_0^{+\infty} \frac{1}{\sqrt{t}} \exp\left(\frac{-x^2}{2t} - t\right) \mathrm{d}t \\
\text{(substitution) } y = \sqrt{t}: &= \frac{1}{\sqrt{2\pi}} \int_0^{+\infty} 2 \exp\left(\frac{-x^2}{2y^2} - y^2\right) \mathrm{d}y \\
\frac{x^2}{2y^2} + y^2 &= \left(y + \frac{|x|}{\sqrt{2}y}\right)^2 - \sqrt{2}|x| \\
z = y + \frac{|x|}{\sqrt{2}y} &: \quad \ldots
\end{aligned}
$$

Example: Gaussian mixture with exponentially distributed variance

$$
\begin{aligned}
f(x) &= \frac{1}{\sqrt{2\pi}} \int_0^{+\infty} \frac{1}{\sqrt{t}} \exp\left(\frac{-x^2}{2t} - t\right) \mathrm{d}t \\
\text{(substitution) } y = \sqrt{t}: &= \frac{1}{\sqrt{2\pi}} \int_0^{+\infty} 2 \exp\left(\frac{-x^2}{2y^2} - y^2\right) \mathrm{d}y \\
\frac{x^2}{2y^2} + y^2 &= \left(y - \frac{|x|}{\sqrt{2}y}\right)^2 + \sqrt{2}|x| \\
z = y - \frac{|x|}{\sqrt{2}y} \quad &: \quad \dots
\end{aligned}
$$

Example: Gaussian mixture with exponentially distributed variance

$$
\begin{aligned}
f(x) &= \frac{1}{\sqrt{2\pi}} \int_0^{+\infty} \frac{1}{\sqrt{t}} \exp\left(\frac{-x^2}{2t} - t\right) \mathrm{d}t \\
\text{(substituition) } y = \sqrt{t}: &= \frac{1}{\sqrt{2\pi}} \int_0^{+\infty} 2 \exp\left(\frac{-x^2}{2y^2} - y^2\right) \mathrm{d}y \\
\text{(substitution) } z &= y - \frac{|x|}{\sqrt{2}y} \\
& \begin{cases} y \to 0 & \Rightarrow z \to -\infty \\ y \to +\infty & \Rightarrow z \to +\infty \end{cases} \\
z^2 &= y^2 + \frac{x^2}{2y^2} - |x|\sqrt{2}
\end{aligned}
$$

Example: Gaussian mixture with exponentially distributed variance

$$
\begin{aligned}
\text{(substitution) } z &= y - \frac{|x|}{\sqrt{2}y} \\
z &= \frac{\sqrt{2}y^2 - |x|}{\sqrt{2}y} \\
y^2 - zy - |x|/\sqrt{2} &= 0 \\
\Delta &= z^2 + 4|x|/\sqrt{2} \\
y &= \frac{z \pm \sqrt{\Delta}}{2} = \frac{z \pm \sqrt{z^2 + 4|x|/\sqrt{2}}}{2} \geq 0 \\
\mathrm{d}y &= \frac{1}{2}\left(\mathrm{d}z + \frac{2z\mathrm{d}z}{2\sqrt{z^2 + 4|x|/\sqrt{2}}}\right)
\end{aligned}
$$

Example: Gaussian mixture with exponentially distributed variance

$$
\begin{aligned}
f(x) &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} \cancel{2} \exp\left(-z^2 - \overbrace{|x|\sqrt{2}}^{\text{out of integral}} \right) \left(\frac{z}{\sqrt{z^2 + 4|x|/\sqrt{2}}} + 1 \right) \\
&= \frac{1}{\sqrt{2\pi}} \exp(-|x|\sqrt{2}) \int_{-\infty}^{+\infty} \exp(-z^2) \left(\frac{z}{\sqrt{z^2 + 4|x|/\sqrt{2}}} + 1 \right) \mathrm{d}z \\
&= \frac{1}{\sqrt{2\pi}} \exp(-|x|\sqrt{2}) \left[\int_{-\infty}^{+\infty} \exp(-z^2) \mathrm{d}z + \int_{-\infty}^{+\infty} \frac{z \exp(-z^2)}{\sqrt{z^2 + 4|x|/\sqrt{2}}} \mathrm{d}z \right.
\end{aligned}
$$

Example: Gaussian mixture with exponentially distributed variance

$$
\begin{aligned}
f(x) &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} \not{2} \exp\left(-z^2 - \overbrace{|x|\sqrt{2}}^{\text{out of integral}}\right)\left(\frac{z}{\sqrt{z^2 + 4|x|/\sqrt{2}}} + 1\right) \\
&= \frac{1}{\sqrt{2\pi}} \exp(-|x|\sqrt{2}) \int_{-\infty}^{+\infty} \exp(-z^2)\left(\frac{z}{\sqrt{z^2 + 4|x|/\sqrt{2}}} + 1\right) \mathrm{d}z \\
&= \frac{1}{\sqrt{2\pi}} \exp(-|x|\sqrt{2}) \left[\underbrace{\int_{-\infty}^{+\infty} \exp(-z^2)\mathrm{d}z}_{=\sqrt{\pi}} + \int_{-\infty}^{+\infty} \frac{z\exp(-z^2)}{\sqrt{z^2 + 4|x|/\sqrt{2}}}\mathrm{d}z \right.
\end{aligned}
$$

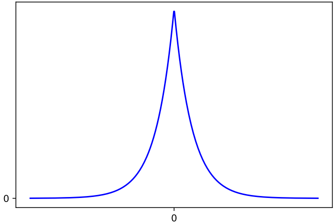
Example: Gaussian mixture with exponentially distributed variance

$$
\begin{aligned}
f(x) &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} \not{2} \exp\left(-z^2 - \overbrace{|x|\sqrt{2}}^{\text{out of integral}}\right) \left(\frac{z}{\sqrt{z^2 + 4|x|/\sqrt{2}}} + 1\right) \\
&= \frac{1}{\sqrt{2\pi}} \exp(-|x|\sqrt{2}) \int_{-\infty}^{+\infty} \exp(-z^2) \left(\frac{z}{\sqrt{z^2 + 4|x|/\sqrt{2}}} + 1\right) \mathrm{d}z \\
&= \frac{1}{\sqrt{2\pi}} \exp(-|x|\sqrt{2}) \left[\underbrace{\int_{-\infty}^{+\infty} \exp(-z^2)\mathrm{d}z}_{=\sqrt{\pi}} + \underbrace{\int_{-\infty}^{+\infty} \frac{z\exp(-z^2)}{\sqrt{z^2 + 4|x|/\sqrt{2}}}\mathrm{d}z}_{\text{odd function of z}} \right.
\end{aligned}
$$

Example: Gaussian mixture with exponentially distributed variance

$$f(x) = \frac{1}{\sqrt{2}} \exp(-|x|\sqrt{2})$$

Laplace Distribution with variance 1



Example: Gaussian mixture with exponentially distributed variance

In practice:

- Histogram of a high pass filtered image might look Laplacian
- In reality it is a Gaussian Mixture with varying σ_i^2
  - Detection in Laplace noise : Delimiter (non linearity) + Correlator
  - Detection in a Gaussian Mixture (known variances): Generalized Matched Filter
- Very different detectors: If we can reliably estimate σ_i^2, GMF will perform better because it uses prior knowledge about the structure