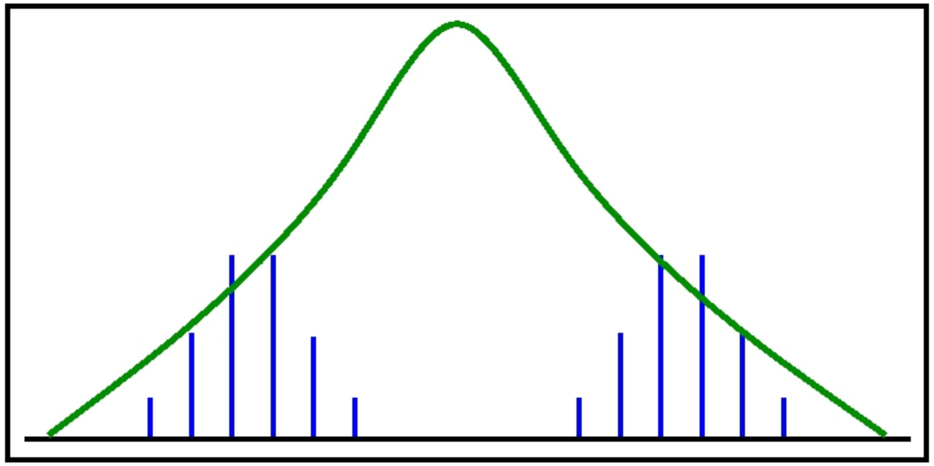
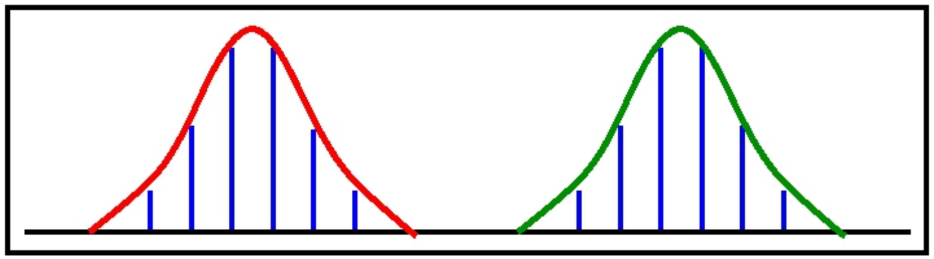
M-of-Gs: Curve-Fitting

- Almost any parameterised curve can be fitted to a bunch of points with an associated error cost
- Gaussian: computationally nice: 2 moments
  - any hump/bump:one Gaussian
  - const region: many closely-spaced Gaussians
- 

  1-G: bad fit as most prob region doesn't have many samples
- 

  2-G: better fit

Home Page

Title Page

Contents

Page 11 of 32

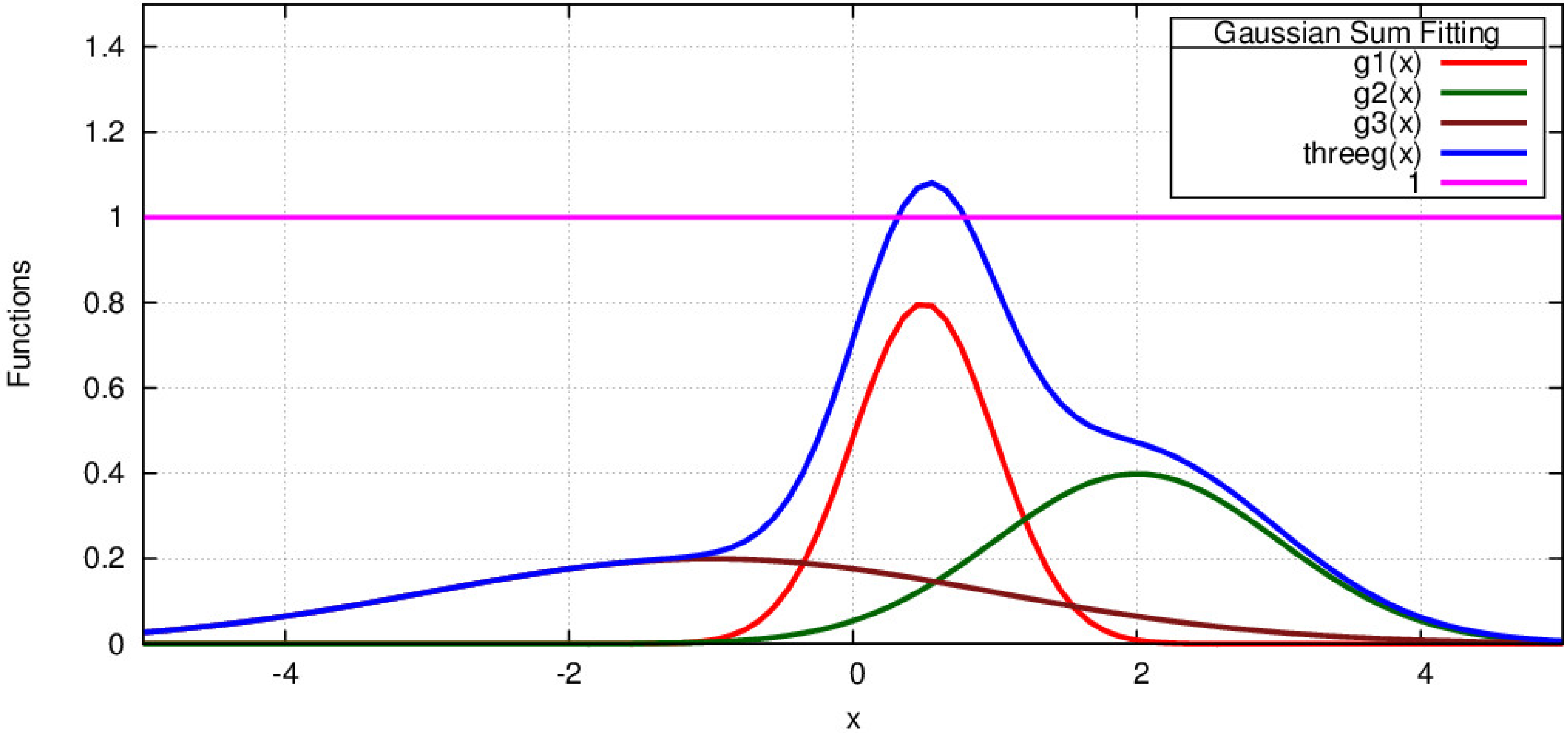
Go Back

Full Screen

Close

Quit

3-G?



- Problem? Probability!
- Solution? Linear combination, prob-scaled sum

$$p(\mathbf{x}) \stackrel{\triangle}{=} \sum_{j=1}^{K} \pi_j \, \mathcal{N}(\mathbf{x} | \boldsymbol{\mu}_j, \boldsymbol{\Sigma}_j)$$

Home Page

Title Page

Contents

Page 12 of 32

Go Back

Full Screen

Close

Quit

Linear Combo: Gaussians

- To finally put it as $\boxed{p(\mathbf{x}) \stackrel{\triangle}{=} \sum_{j=1}^{K} \pi_j \, \mathcal{N}(\mathbf{x}|\boldsymbol{\mu}_j, \boldsymbol{\Sigma}_j)}$
- How? Factorise the marginal
- $p(\mathbf{x}) = \sum_{\forall j} p(\mathbf{x}|j)p(j) = \sum_{j=1}^{K} p(j)p(\mathbf{x}|j)$ Compare!
- π_j: prior prob of picking the jth component
- $p(\mathbf{x}|j) = \mathcal{N}(\mathbf{x}|\boldsymbol{\mu}_j, \boldsymbol{\Sigma}_j)$
- What about the posterior probability? $p(j|\mathbf{x})$
- $\gamma_j(\mathbf{x}) \stackrel{\triangle}{=} p(j|\mathbf{x})$ = 'Responsibility': How much is j responsible for the $\mathbf{x}$, given that $\mathbf{x}$ is observed?
- $p(j|\mathbf{x}) = \frac{p(\mathbf{x}|j)p(j)}{\sum_{\forall l} p(\mathbf{x}|l)p(l)} = \frac{p(j)p(\mathbf{x}|j)}{\sum_l p(l)p(\mathbf{x}|l)}$
- Responsibility $\boxed{\gamma_j(\mathbf{x}) \stackrel{\triangle}{=} \frac{\pi_j \, \mathcal{N}(\mathbf{x}|\boldsymbol{\mu}_j, \boldsymbol{\Sigma}_j)}{\sum_l \pi_l \, \mathcal{N}(\mathbf{x}|\boldsymbol{\mu}_l, \boldsymbol{\Sigma}_l)}}$

Men of God. . .

Thomas Bayes [1701-1761] | G. J. Mendel [1882-1884] | M. Mitra [1968-]

https://upload.wikimedia.org/wikipedia/commons/d/d4/Thomas_Bayes.gif

https://upload.wikimedia.org/wikipedia/commons/3/3d/Gregor_Mendel_oval.jpg

http://iseeindia.com/wordpress/wp-content/uploads/2011/11/Ramkrishna_Miss11736-290x290.jpg

Mahan Maharaj/Swami Vidyanathananda
2011 Shanti Swarup Bhatnagar Award in Math Sciences
Infosys Prize 2015 for Mathematical Sciences

Aside: Life of π_j: Properties

[http://i1.ytimg.com/vi/j9Hjrs6WQ8M/maxresdefault.jpg] Richard Parker

- $\pi_j \in [0, 1]$
- $\sum_{j=1}^{K} \pi_j = 1$
- How, and Why?

$$p(\mathbf{x}) \stackrel{\triangle}{=} \sum_{j=1}^{K} \pi_j \; \mathcal{N}(\mathbf{x} | \boldsymbol{\mu}_j, \boldsymbol{\Sigma}_j)$$

$p, \mathcal{N}$: pdfs; $\int_{-\infty}^{+\infty}$ pdf$=1$

Home Page

Title Page

Contents

Page 15 of 32

Go Back

Full Screen

Close

Quit

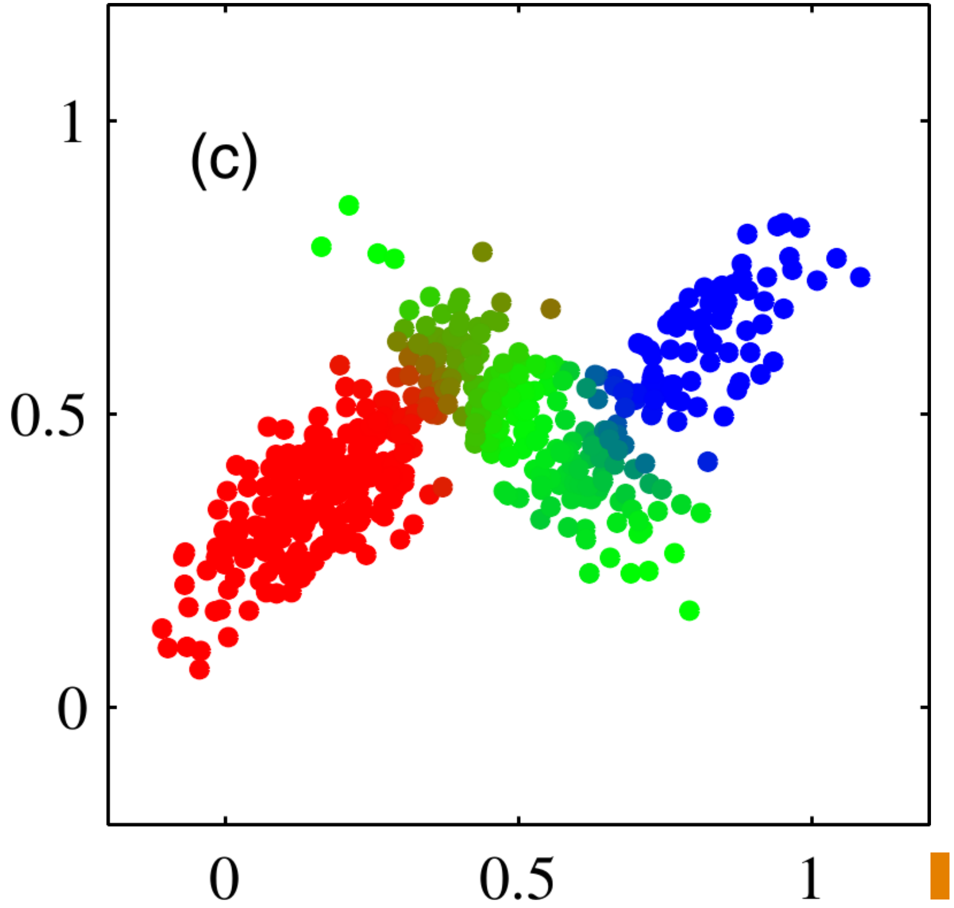
1. 500 points from a mixture of 3 Gaussians

   (a) Samples, the source specified: joint $p(j)p(\mathbf{x}_i|j)$

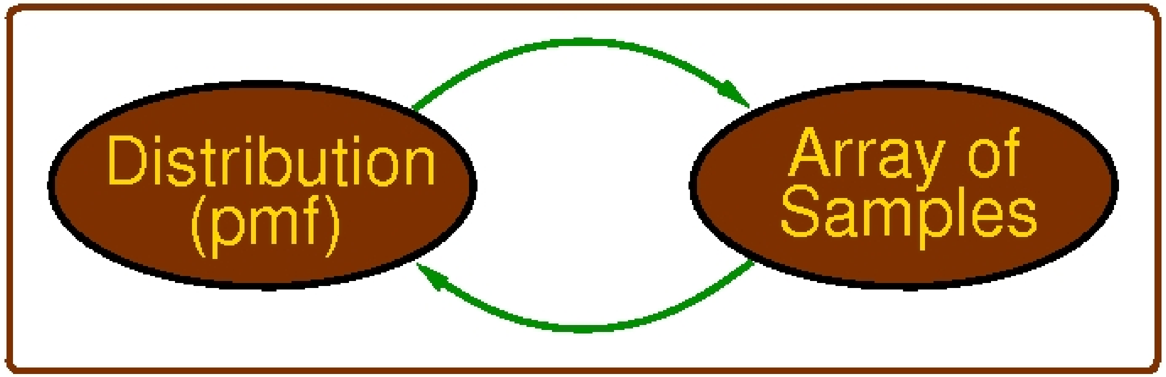
   (b) Just the marginal $p(\mathbf{x}_i)$, ignoring j

   (c) responsibilities $\gamma_j(\mathbf{x}_i)$ (RGB proportion) for $\mathbf{x}_i$

[C. M. Bishop, Pattern Recognition and Machine Learning]



Given a set of samples



- Generating samples acc to a distribution: uniform, Gaussian (Box-Müller), general ($y-$ axis)
- Ideal case: The density is indeed a mixture of K Gaussians, as modelled. Draw a set of N samples
- Actual/Reality Just given N observations $\{\mathbf{x}_i\}$ from a physical process. Assume a model: mixture of K Gaussians, to estimate its parameters

Param Estimation: 1 1-D Gaussian

- Given: $X = \{x_1, \ldots x_N\}$, a set of N observations
- Model: One 1-D Gaussian, mean μ, variance σ^2
- Assumptions: Data points i.i.d. Independent: allows marginal prob multiplication without considering conditional dependence terms. Identically distributed: all from same model
- Method: $p(X|\ \mu,\ \sigma^2)$: Likelihood, to maximise
- Reasonable? Find params which maximise the likelihood of getting these points, given our model
- $p(X|\ \mu,\ \sigma^2) = \prod_{i=1}^{N} \mathcal{N}(x_i|\ \mu,\ \sigma^2)$, to maximise
- $\equiv$ Maximise log-likelihood. Why? (John Napier)
  - increasing function $\implies$ same nature
  - Multiplications $\rightarrow$ additions