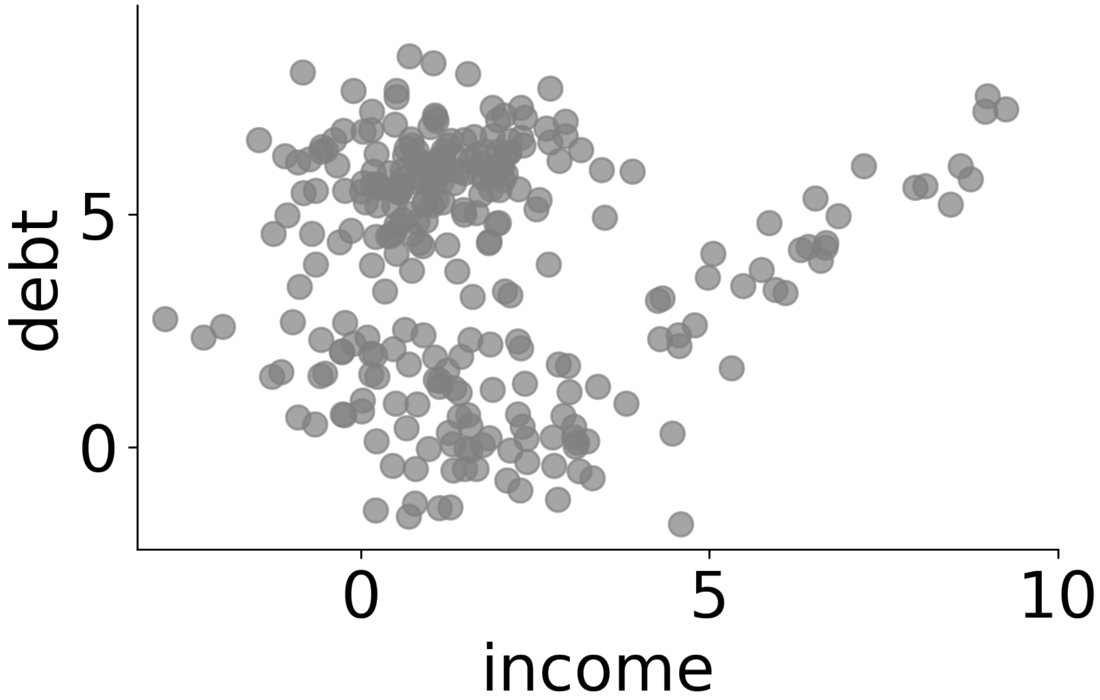
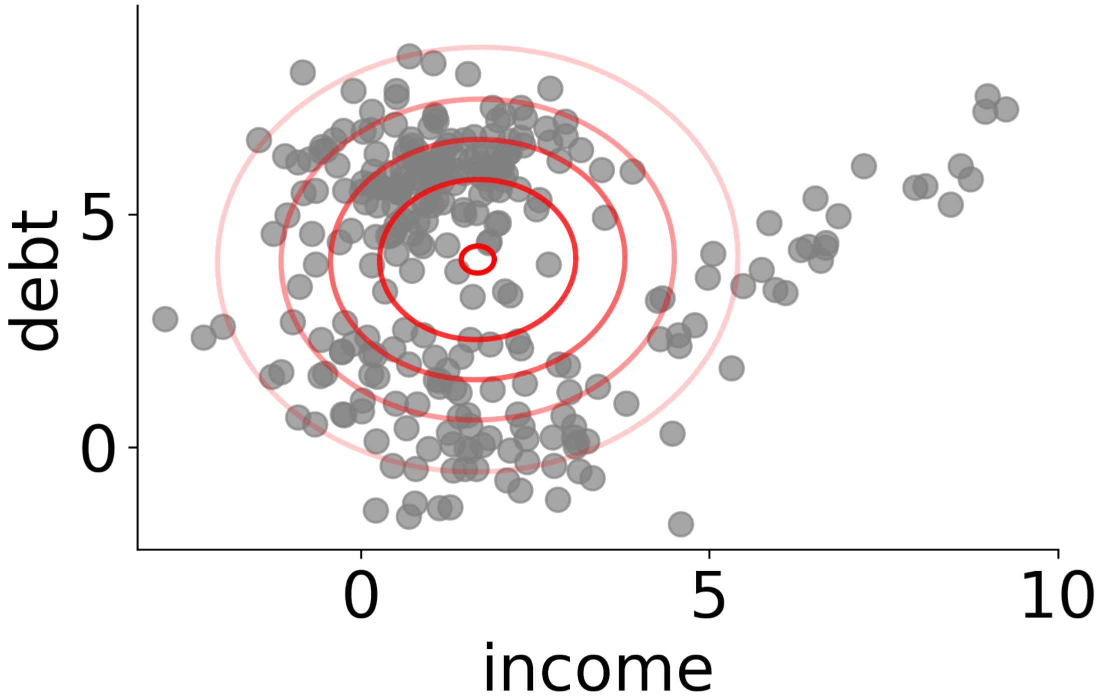
Probabilistic model of data



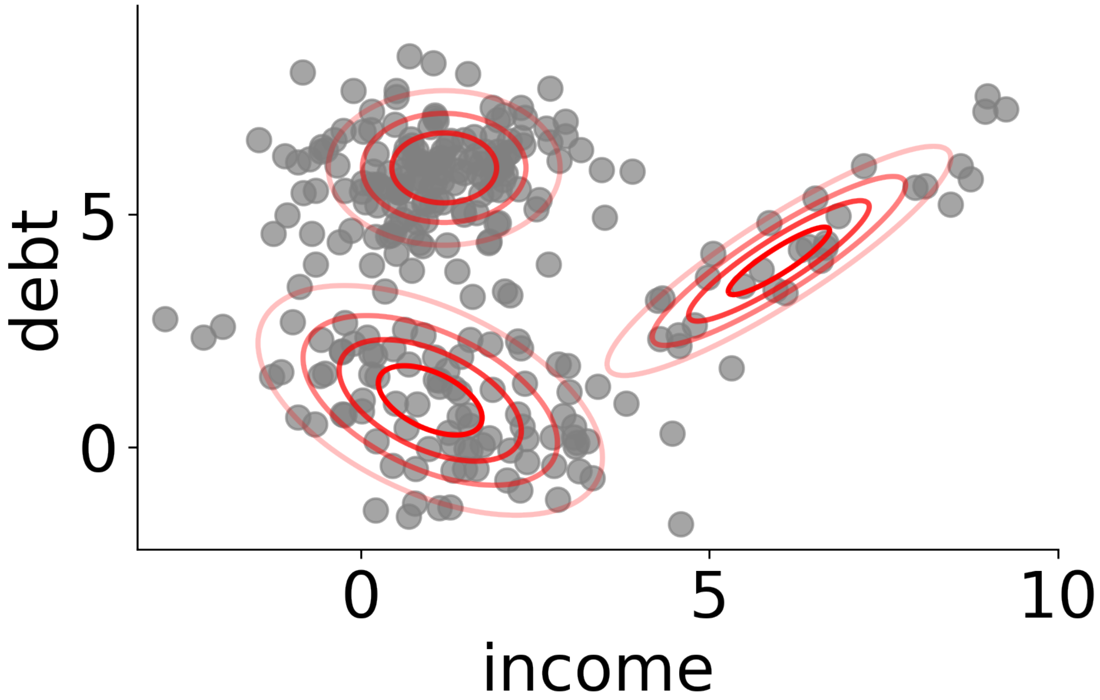
Probabilistic model of data



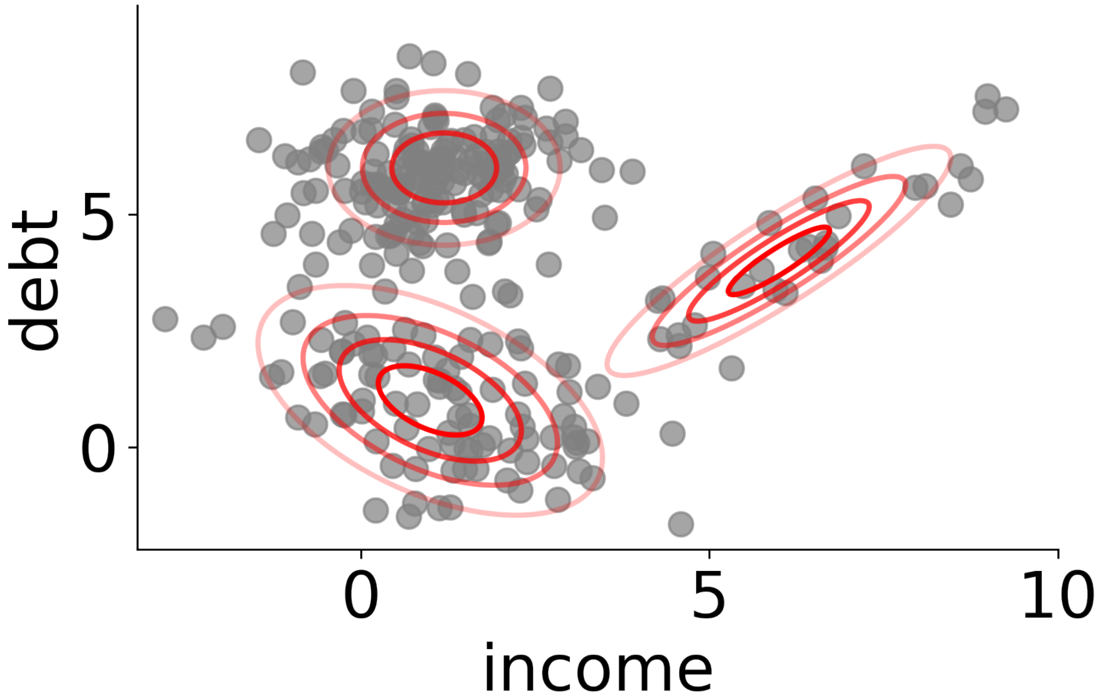
$$p(x \mid \theta) = \mathcal{N}(x \mid \mu, \Sigma)$$

$$\theta = \{\mu, \Sigma\}$$

Gaussian Mixture Model (GMM)

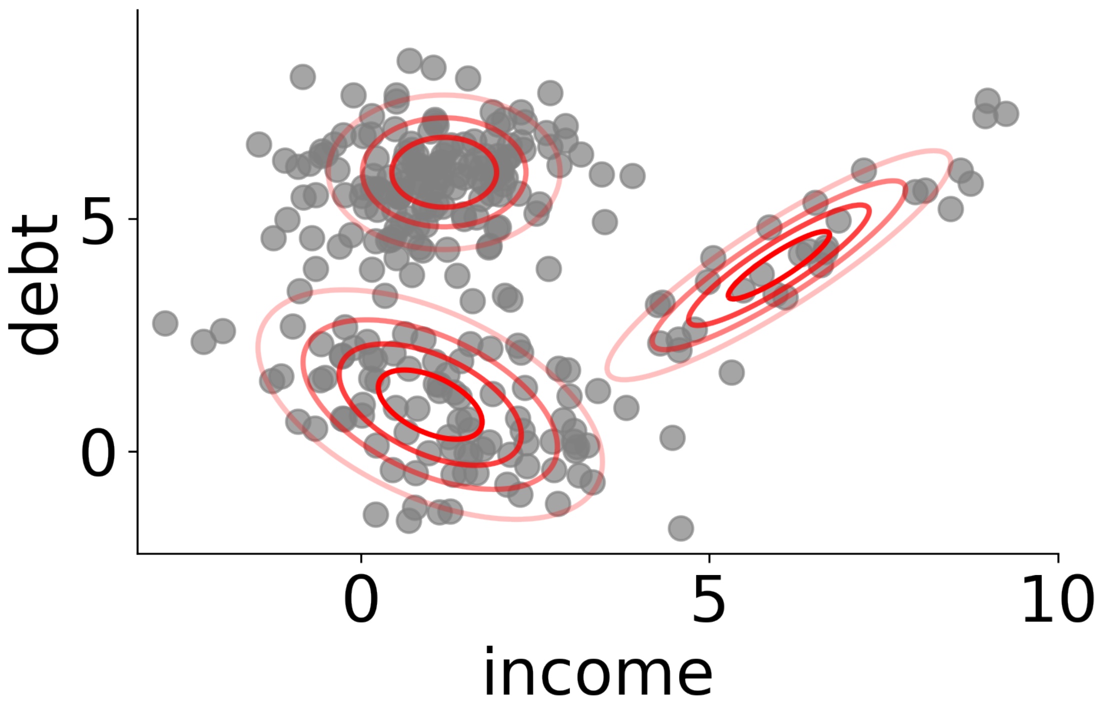


Gaussian Mixture Model (GMM)



$$p(x \mid \theta) = \pi_1 \mathcal{N}(x|\mu_1, \Sigma_1) + \pi_2 \mathcal{N}(x|\mu_2, \Sigma_2) + \pi_3 \mathcal{N}(x|\mu_3, \Sigma_3)$$

Gaussian Mixture Model (GMM)



$$p(x \mid \theta) = \pi_1 \mathcal{N}(x|\mu_1, \Sigma_1) + \pi_2 \mathcal{N}(x|\mu_2, \Sigma_2) + \pi_3 \mathcal{N}(x|\mu_3, \Sigma_3)$$

$$\theta = \{\pi_1, \pi_2, \pi_3, \mu_1, \mu_2, \mu_3, \Sigma_1, \Sigma_2, \Sigma_3\}$$

Gaussian Mixture Model (GMM)

	Gaussian	GMM
Flexibility		

GMM vs Guassian

	Gaussian	GMM
Flexibility		
# of parameters		
Parameters	μ, Σ	$\{\pi_1, \pi_2, \pi_3\}$ $\{\mu_1, \mu_2, \mu_3\}$ $\{\Sigma_1, \Sigma_2, \Sigma_3\}$

Training GMM

$$\max_{\theta}. \quad p(X \mid \theta)$$

Training GMM

$$\max_{\theta} \quad p(X \mid \theta) = \prod_{i=1}^{N} p(x_i \mid \theta)$$

Training GMM

$$\max_{\theta}. \quad \prod_{i=1}^{N} p(x_i \mid \theta) = \prod_{i=1}^{N} \left(\pi_1 \mathcal{N}(x_i \mid \mu_1, \Sigma_1) + \ldots\right)$$

Training GMM

$$\max_{\theta} \quad \prod_{i=1}^{N} p(x_i \mid \theta) = \prod_{i=1}^{N} \left(\pi_1 \mathcal{N}(x_i \mid \mu_1, \Sigma_1) + \ldots\right)$$

$$\text{subject to} \quad \pi_1 + \pi_2 + \pi_3 = 1;\ \pi_k \geq 0;\ k = 1, 2, 3.$$

Training GMM

$$\max_{\theta} \quad \prod_{i=1}^{N} p(x_i \mid \theta) = \prod_{i=1}^{N} \left(\pi_1 \mathcal{N}(x_i \mid \mu_1, \Sigma_1) + \ldots\right)$$

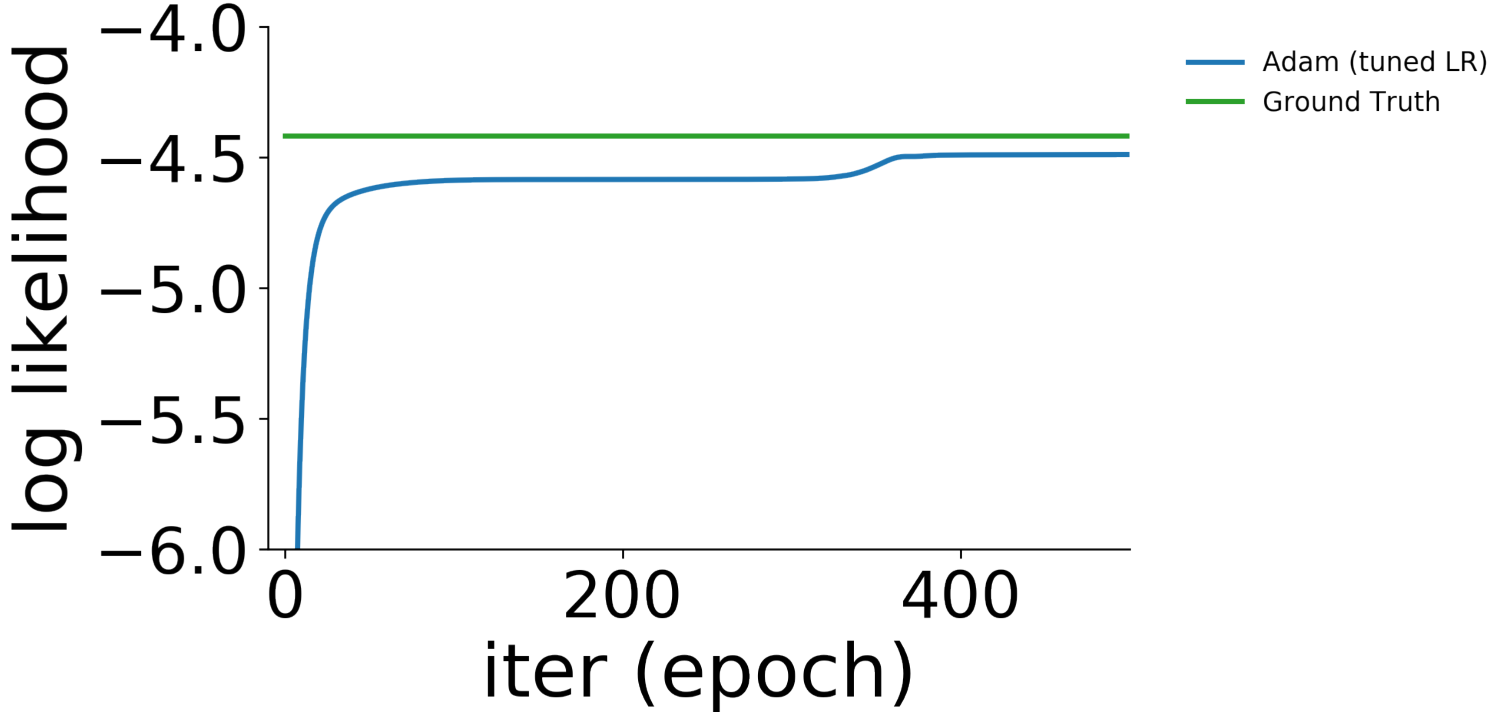
$$\text{subject to} \quad \pi_1 + \pi_2 + \pi_3 = 1;\ \pi_k \geq 0;\ k = 1, 2, 3.$$

$$\Sigma_k \succ 0;$$

Training GMM

$$\max_{\theta} \quad \prod_{i=1}^{N} p(x_i \mid \theta) = \prod_{i=1}^{N} \left(\pi_1 \mathcal{N}(x_i \mid \mu_1, \Sigma_1) + \ldots\right)$$

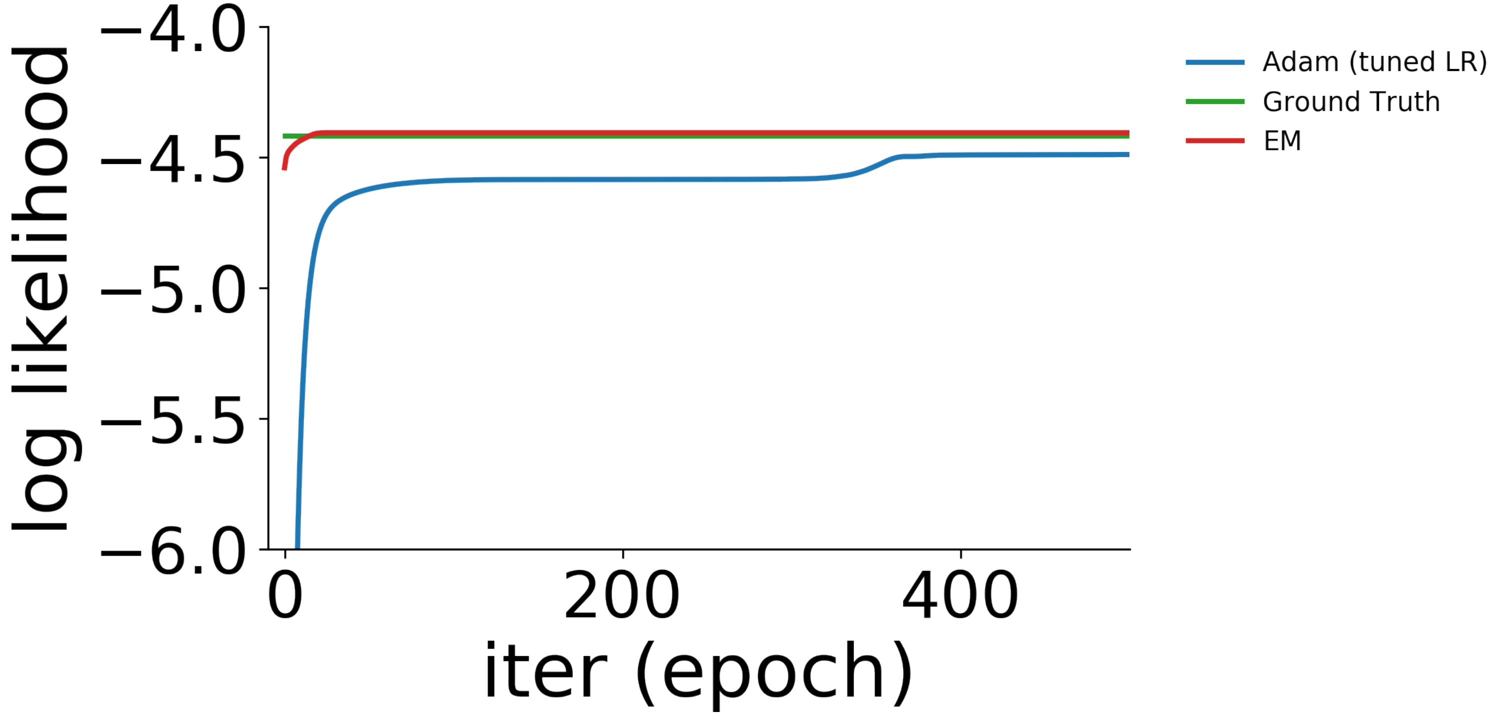
$$\text{subject to} \quad \pi_1 + \pi_2 + \pi_3 = 1;\ \pi_k \geq 0;\ k = 1, 2, 3.$$



Training GMM

$$\max_{\theta}. \quad \prod_{i=1}^{N} p(x_i \mid \theta) = \prod_{i=1}^{N} \left(\pi_1 \mathcal{N}(x_i \mid \mu_1, \Sigma_1) + \ldots\right)$$

$$\text{subject to} \quad \pi_1 + \pi_2 + \pi_3 = 1;\ \pi_k \geq 0;\ k = 1, 2, 3.$$



Summary

- Gaussian Mixture Model is a flexible probability distribution
- It is hard to fit (train) with SGD