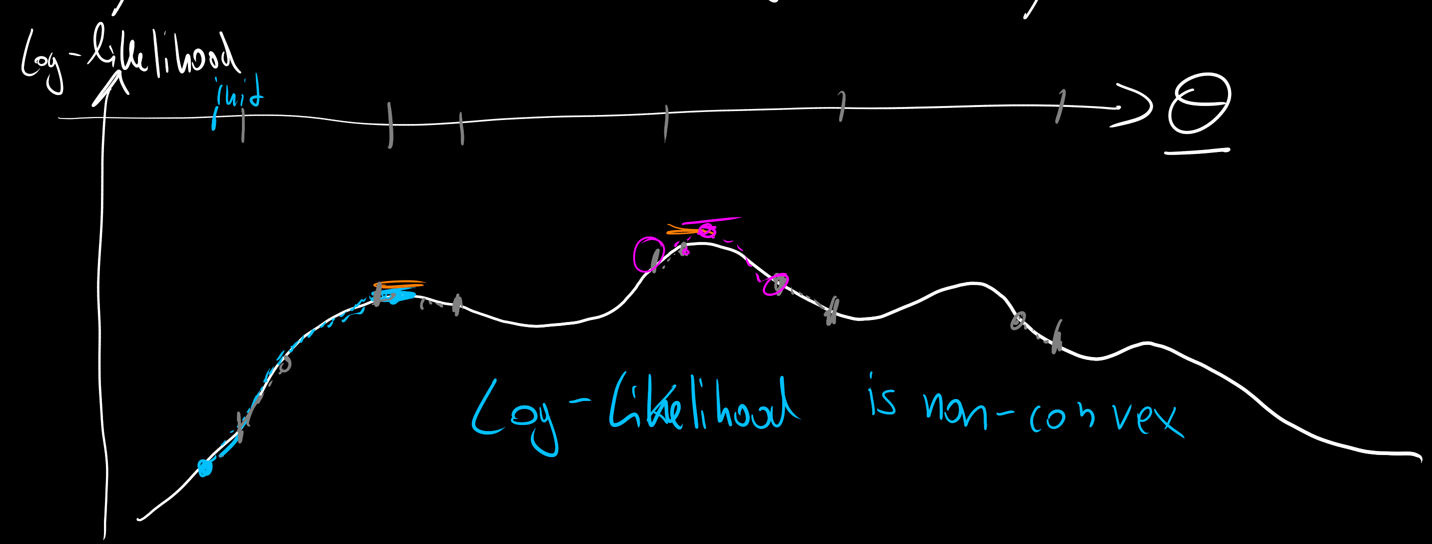
Maximizing the marginal log-likelihood

$$\underline{\pi}, \underline{\mu}, \underline{\Sigma} = \arg\max \log p(D; \underline{\pi}, \underline{\mu}, \underline{\Sigma})$$



Marginal Likelihood

$$p(D; \ldots) = \prod_{i=0}^{N-1} p(X = x^{[i]})$$

Marginalize out the latent Z

$$\mathcal{L}(D) = \prod_{i=0}^{N-1} \sum_{d=0}^{D-1} p(Z=d, X=x^{[i]})$$

$$= \prod_{i=0}^{N-1} \sum_{d=0}^{D-1} p(Z=d)\, p(X=x^{[i]} \mid Z=d)$$

$$= \prod_{i=0}^{N-1} \sum_{d=0}^{D-1} \pi_d \, \mathcal{N}(X=x^{[i]}, \mu_d, \Sigma_d)$$

Log-Likelihood

$$\ell(D) = \log(\mathcal{L}(D))$$

$$= \sum_{i=0}^{N-1} \log \sum_{d=0}^{D-1} \pi_d \, \mathcal{N}(X=x^{[i]}; \mu_d, \Sigma_d)$$

$$= \sum_{i=0}^{N-1} \underbrace{\log \sum_{d=0}^{D-1} \exp}_{\underset{d=0}{\overset{D-1}{\text{logsumexp}}}} \left(\log \pi_d + \log \mathcal{N}(X=x^{[i]}; \mu_d, \Sigma_d) \right)$$

So far

initialize π, μ, Σ
for $i = 0:100$
    E-Step
    M-Step
end

Now

initialize π, μ, Σ
for $i = 0:100$
    E-Step
    M-Step
    calculate log-likelihood
end

Sieving

① Pre-Sieving
for 100 candidates
    initialize π, μ, Σ
    for $i = 0:5$
        E-Step
        M-Step
        calculate log-likelihood
    end
end

② Sieving
select 10 candidates based on highest likelihood

③ Post-Sieving
for the 10 chosen ones
    for $i = 0:100$
        E-Step
        M-Step
        calculate log-likelihood
    end
end

④ Select the best