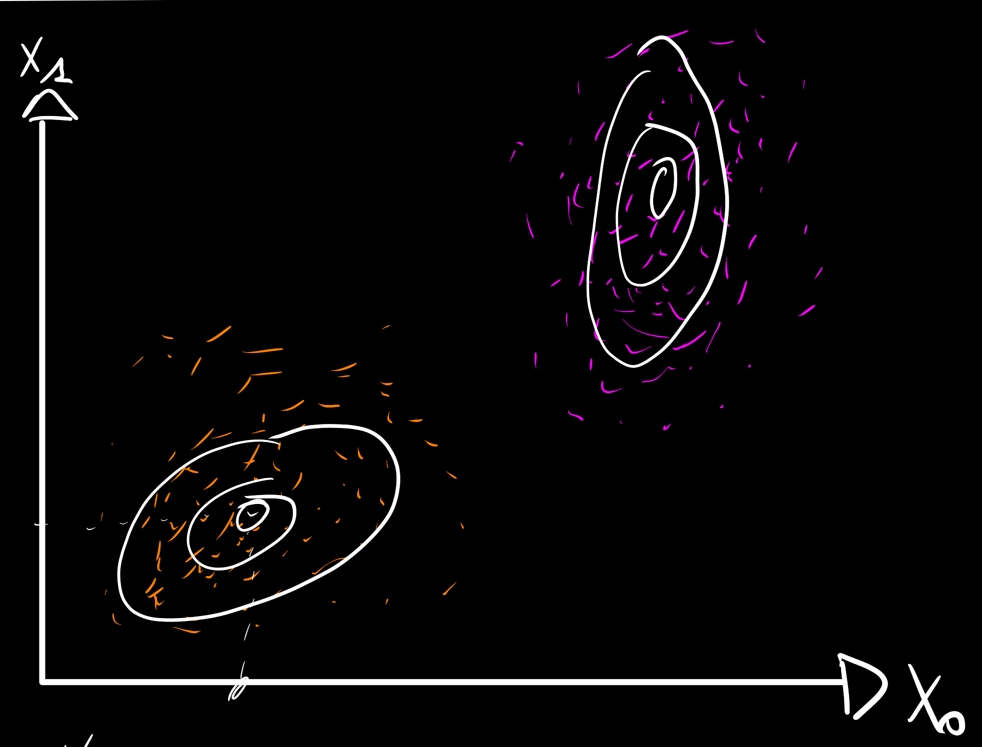
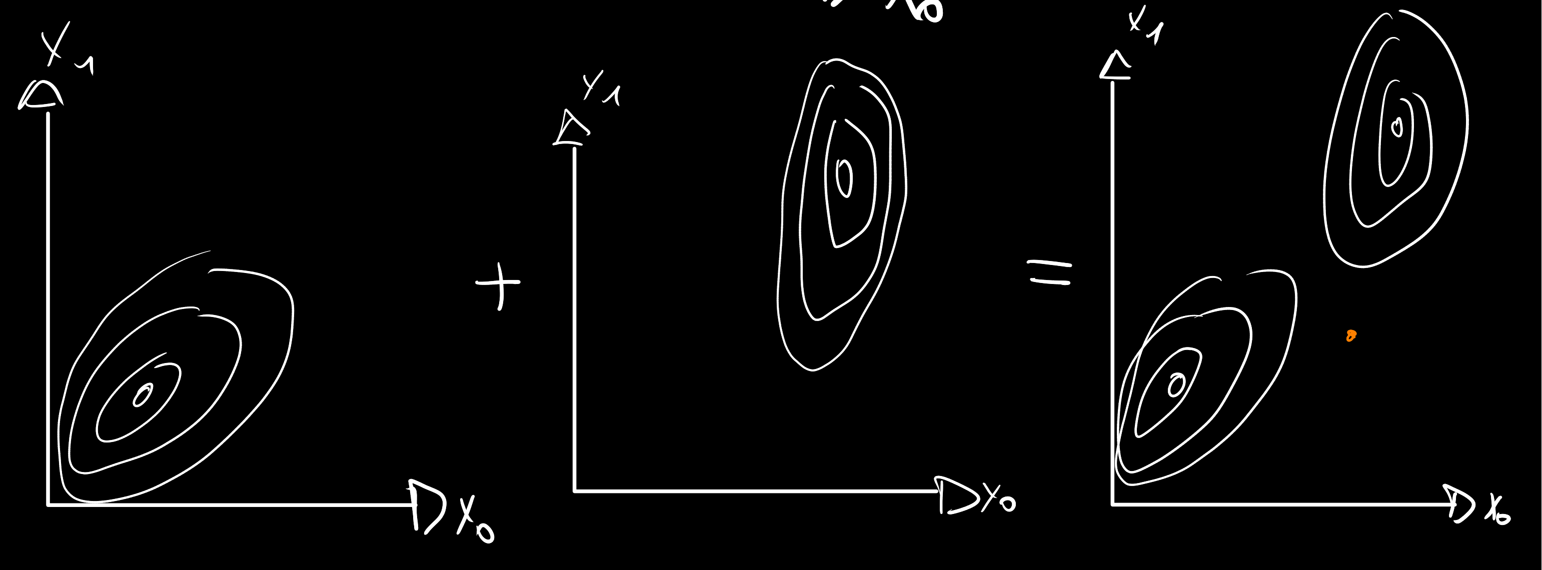
Multivariate Gaussian Mixture Models





$$p_0(\underline{x}) + p_1(\underline{x}) \neq p(\underline{x})$$

violates normalization

↳ introduce "weighting" coefficients

$$p(\underline{x}) = \pi_0 p_0(\underline{x}) + \pi_1 p_1(\underline{x})$$

$$\underset{\text{in general}}{=} \sum_{d=0}^{D-1} \pi_d \, p_d(\underline{x})$$

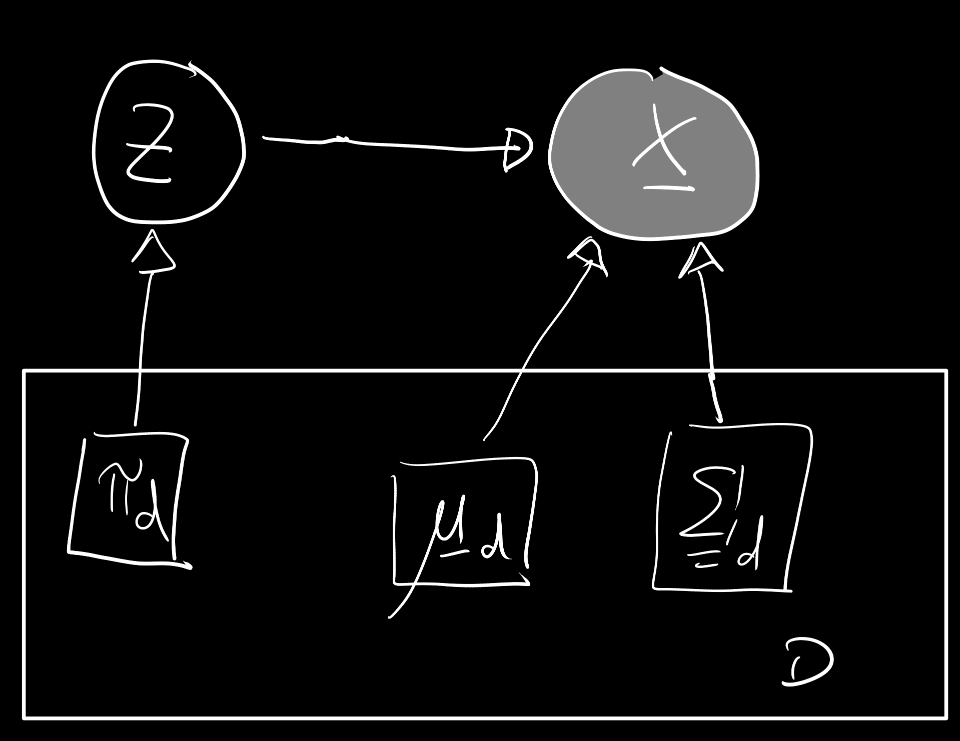
d Multivariate Normals

⇒ this is the marginal $p(\underline{x})$

↳ the assignment to one of the clusters/classes is a latent information

Z ... cluster assignment

$$Z \in \{0, \dots, D-1\}$$



$\underline{X} \sim$ Multivariate Normal

$Z \sim$ Categorical

$Z \sim Cat(Z; \underline{\pi})$

$\underline{X} \sim \mathcal{N}(\underline{X}; \underline{\mu}_Z, \underline{\underline{\Sigma}}_Z)$

D ... number of clusters/classes

joint:

$$p(Z, \underline{X}) = p(Z)\, p(\underline{X} \mid Z)$$

$$= Cat(Z; \underline{\pi})\, \mathcal{N}(\underline{X}; \underline{\mu}_Z, \underline{\underline{\Sigma}}_Z)$$

$$= \left(\prod_{d=0}^{D-1} \pi_d^{\mathbb{I}(Z=d)} \right) \frac{1}{\sqrt{(2\pi)^K \det(\underline{\underline{\Sigma}}_Z)}} \exp\left(-\frac{1}{2} (\underline{X} - \underline{\mu}_Z)^T \underline{\underline{\Sigma}}^{-1} (\underline{X} - \underline{\mu}_Z)\right)$$

Marginalization

$$p(\underline{X}) = \sum_{d=0}^{D-1} p(Z=d, \underline{X})$$

\# parameters

Categorical	Normal	
D	$D \cdot K$	$D \cdot \frac{K(K+1)}{2}$

$$\underbrace{D \cdot \left(1 + K + \frac{K(K+1)}{2}\right)}_{}$$

$\mathcal{O}(K^2)$

problematic in high-dim spaces

‼ 3 forms of Multivariate Normals / Gaussian

	separate	shared
full	$D\left(1+K+\frac{K(K+1)}{2}\right)$	$D(1+K) + \frac{K(K+1)}{2}$
diagonal	$D(1+K+K)$	$D(1+K) + K$
isotropical / spherical	$D(1+K+1)$	$D(1+K) + 1$

e.g. all have variance of $\sigma^2 = 1$

↳ we always have to save $\underline{\pi} \in \mathbb{R}^D$ and D times $\underline{\mu}_d \in \mathbb{R}^K$

↳ savings occur in the covariance matrix