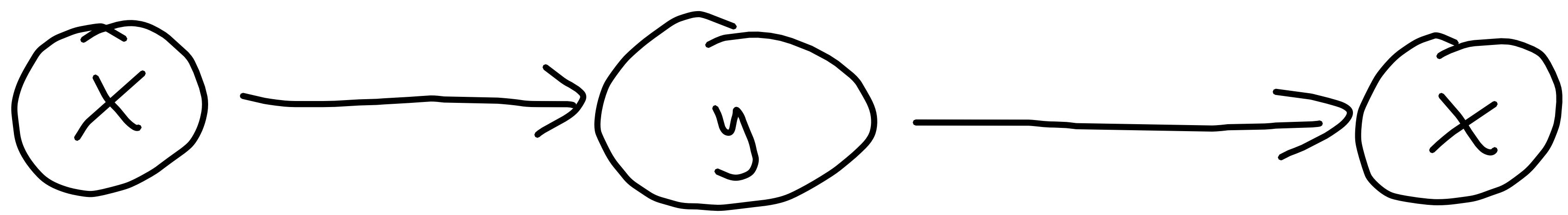
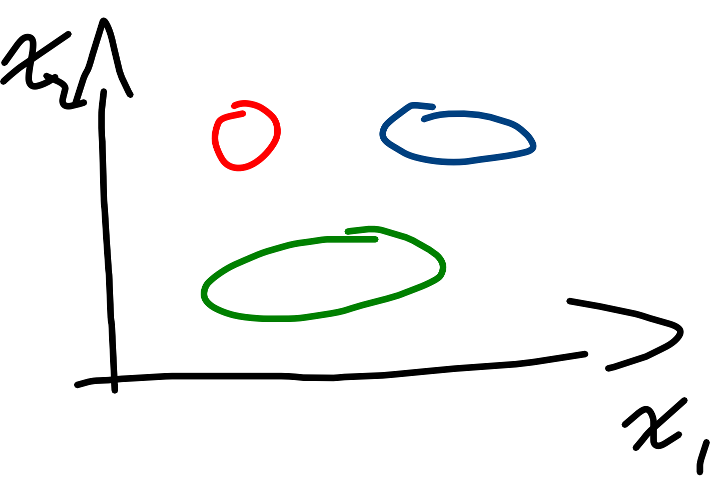
Gaussian Mixture Clustering: EM

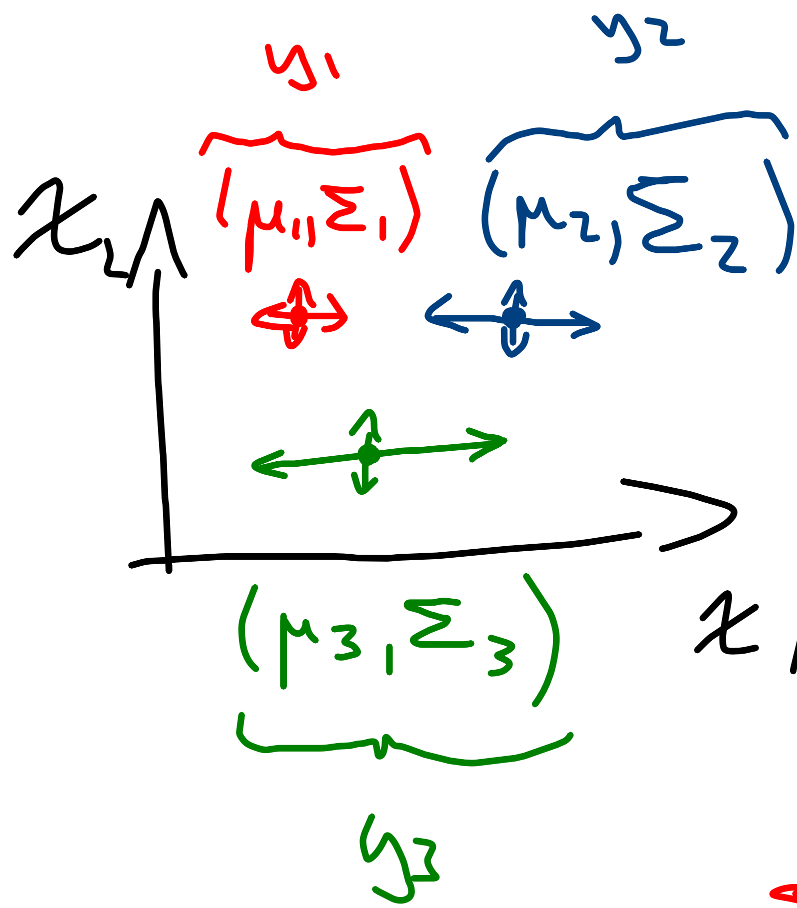
$p_\phi(y|x)$ $\quad p(y) \quad$ $p_\theta(x|y)$

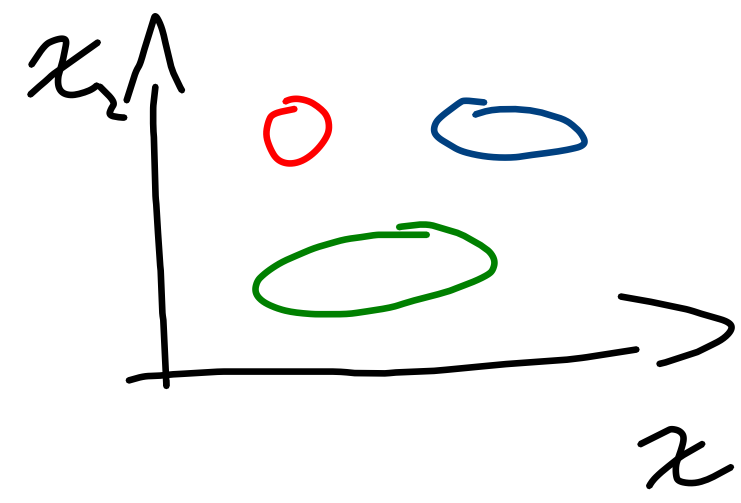
Gaussian that depends on y

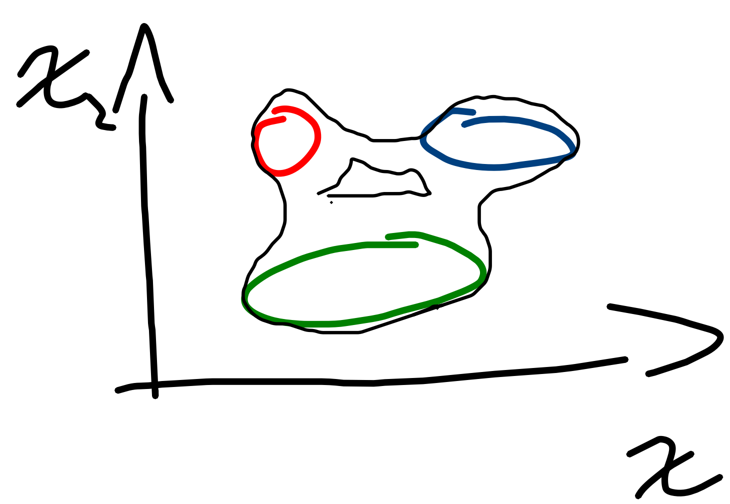


$$p(x) = \int p(x|z) \cdot p(z)\, dz$$

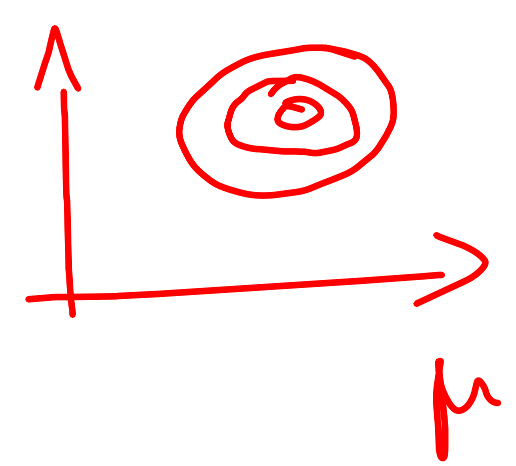


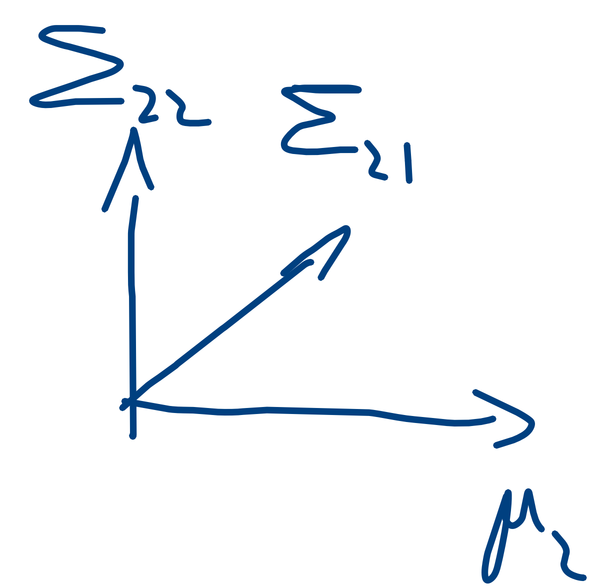






$\Sigma_1 = \sigma I$

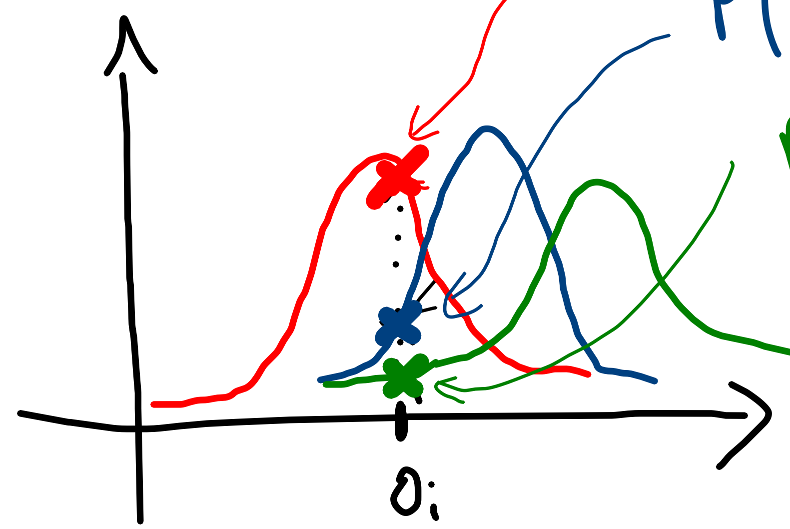




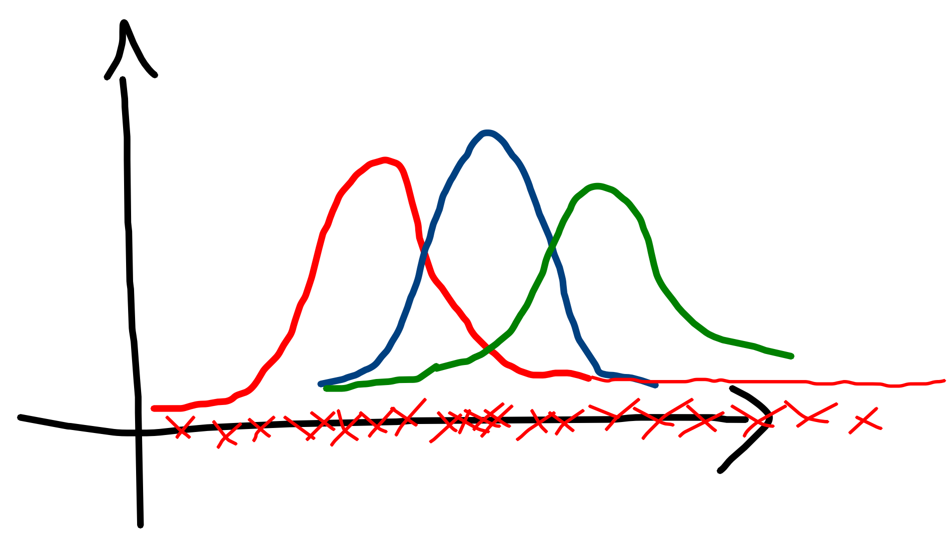
$$P(o_i|k_r) = N(o_i|\mu_r, \Sigma_r)$$

$$P(o_i|k_b) = N(o_i|\mu_b, \Sigma_b)$$

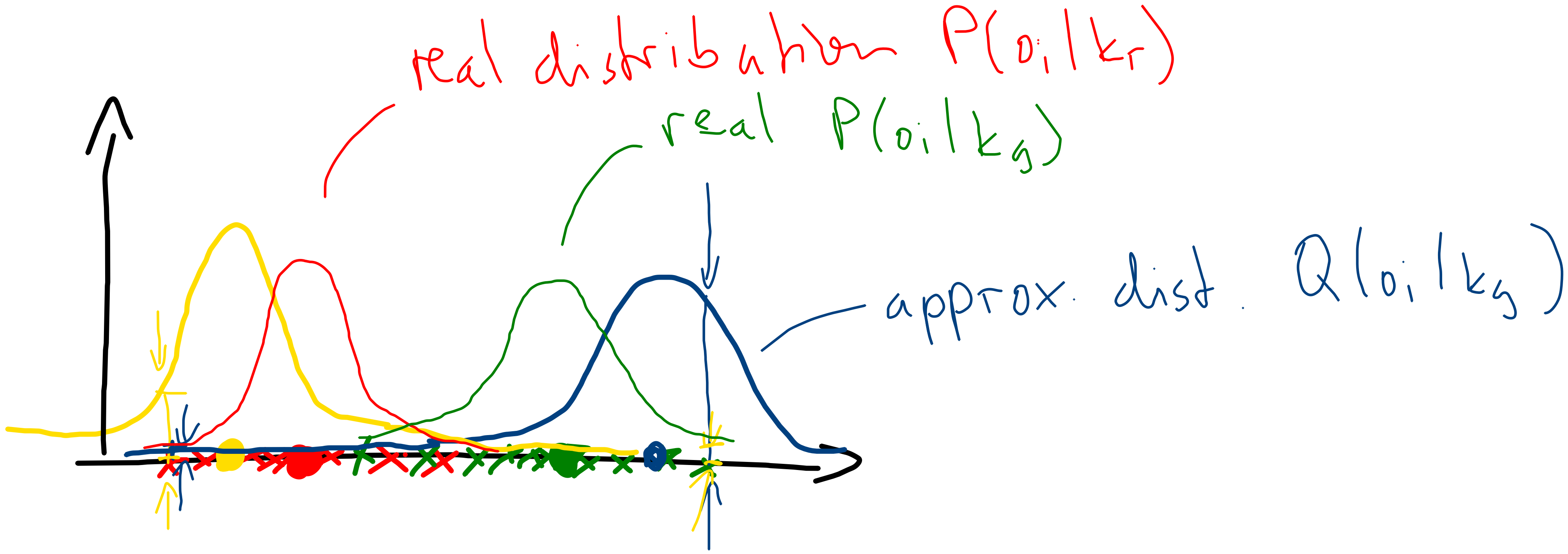
$$P(o_i|k_g) = N(o_i|\mu_g, \Sigma_g)$$



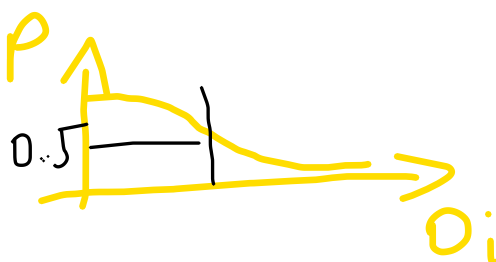
$$P_\theta(k_r|o_i) = \frac{P_\theta(o_i|k_r)\cdot P(k_r)}{P_\theta(o_i|k_r)\cdot P(k_r) + P_\theta(o_i|k_b)\cdot P(k_b) + P_\theta(o_i|k_g)\,P(k_g)}$$



$$\mu_r = \frac{1}{\sum_o P_\theta(k_r|o_i)} \sum_o P_\theta(k_r|o_i)\, o$$

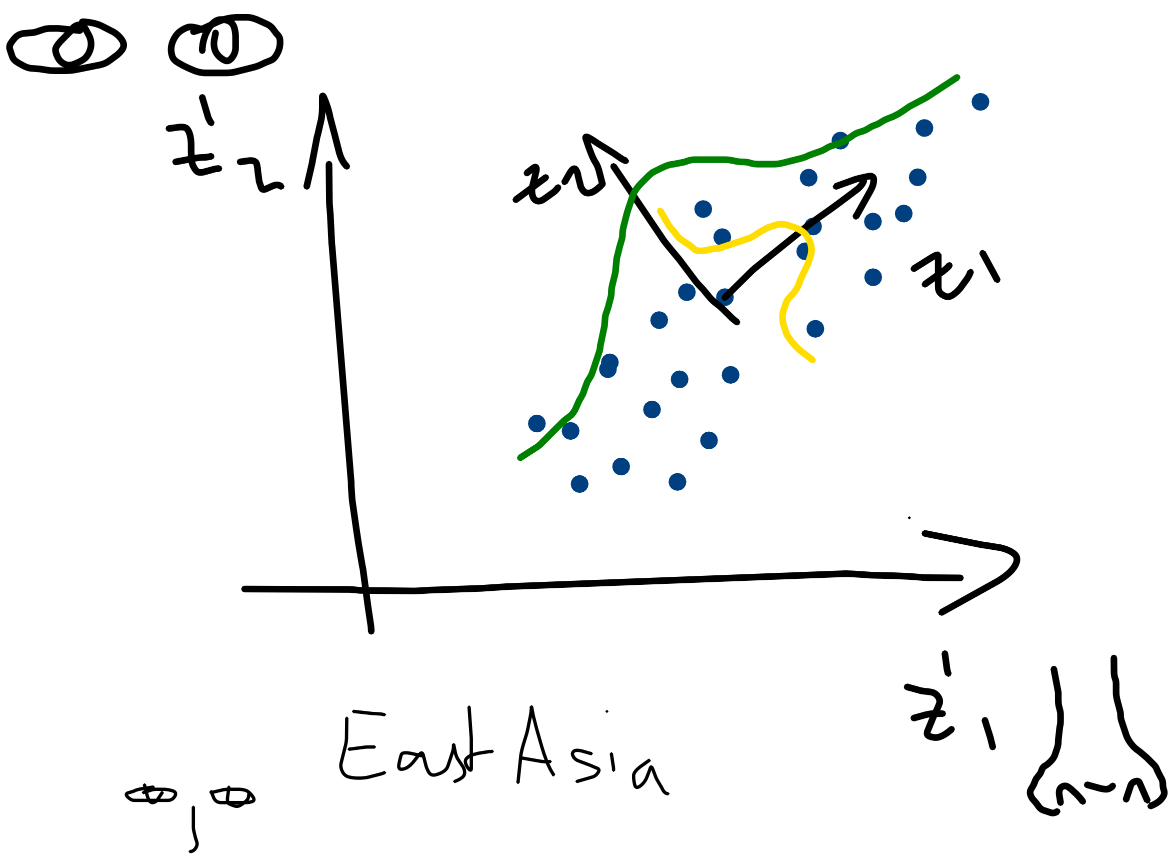


$$Q(k_r \mid o_i) = \frac{Q(o_i \mid k_r) \cdot P(k_r)}{Q(o_i \mid k_r) \cdot P(k_r) + Q(o_i \mid k_g)\, P(k_g)}$$



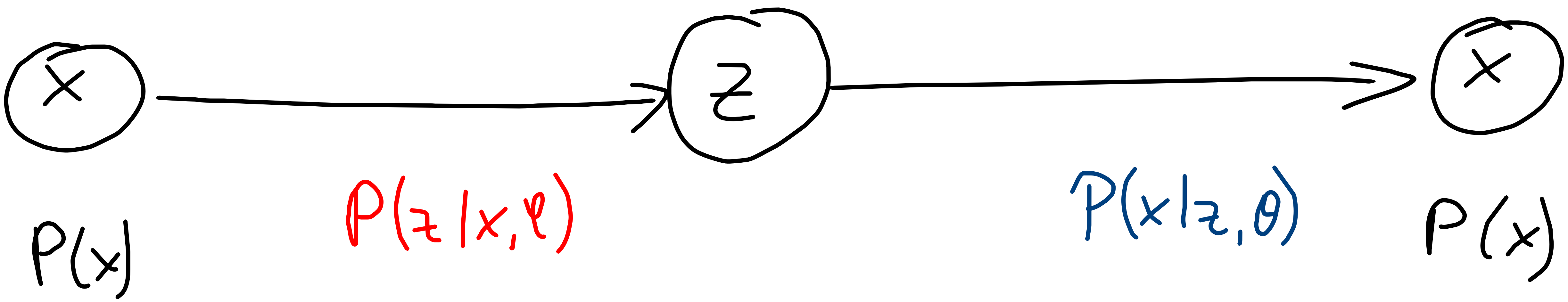
$$\mu_r^{l+1} = \frac{1}{\sum_i Q(k_r \mid o_i)} \sum_i Q(k_r \mid o_i) \cdot o_i$$

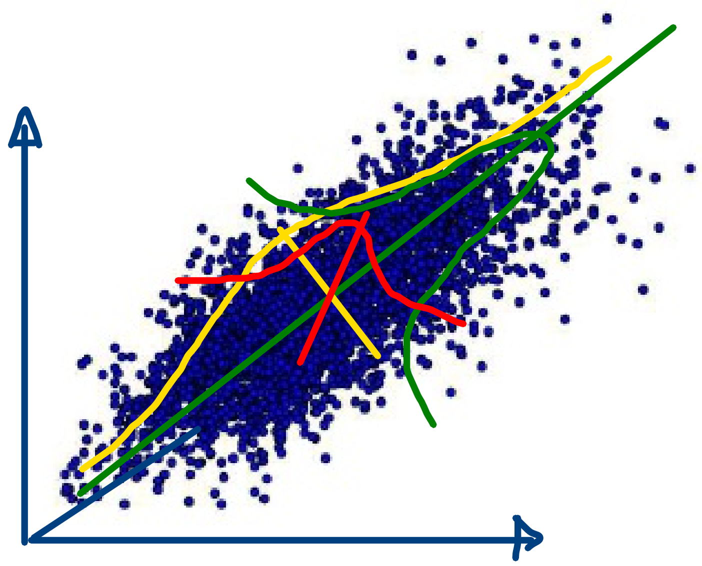
Linear Gaussian Model of Faces

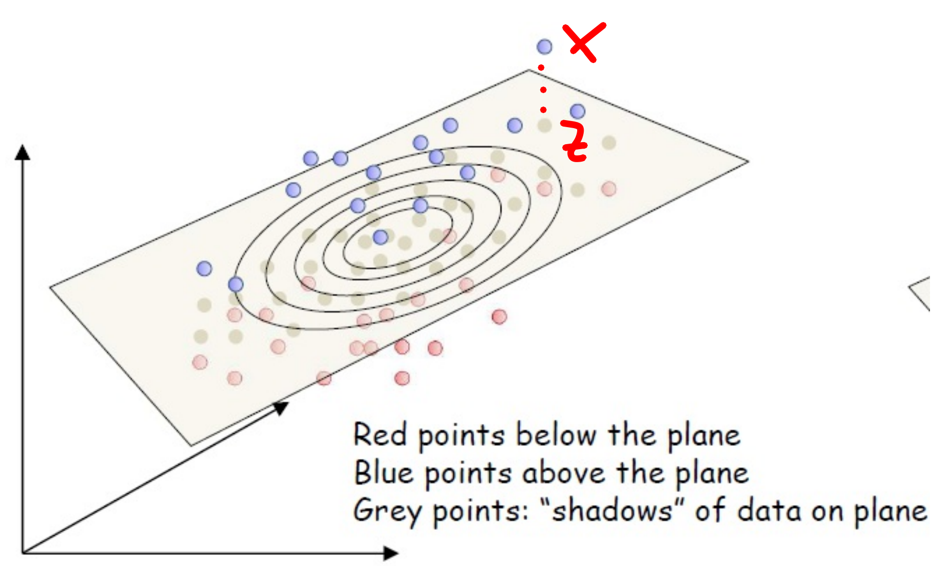


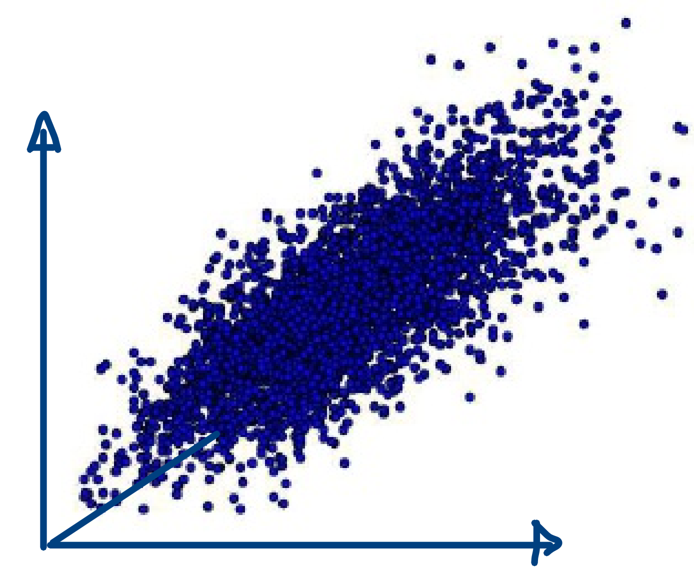
"Z's" are Eigenvalues of the data

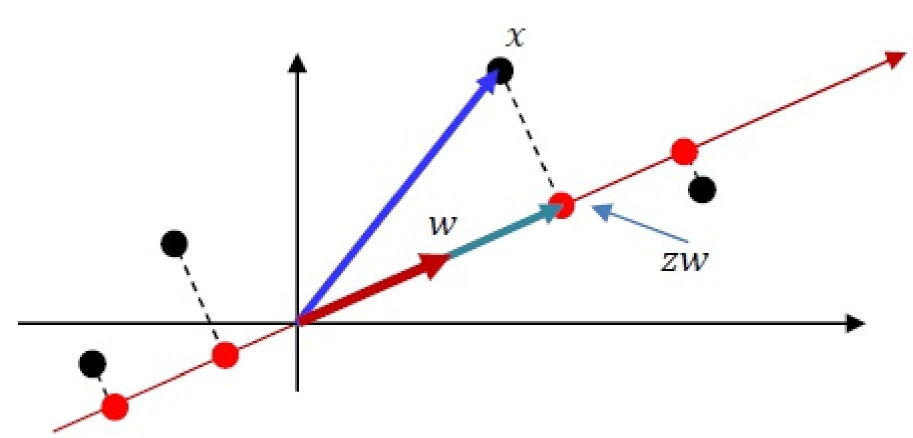
VAE Intuition: PCA aka. Instance of LGM











$w = \varphi$ → Eigenvector

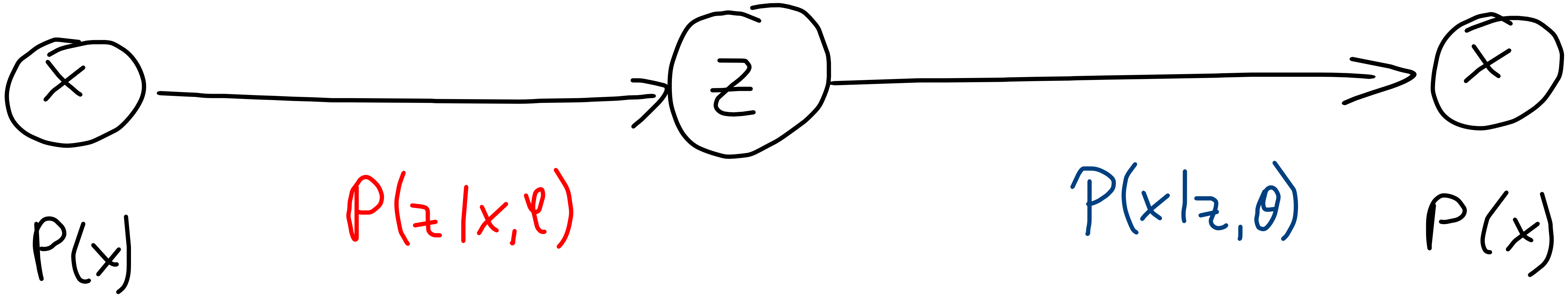
z → Eigenvalue

zw → Projection of x

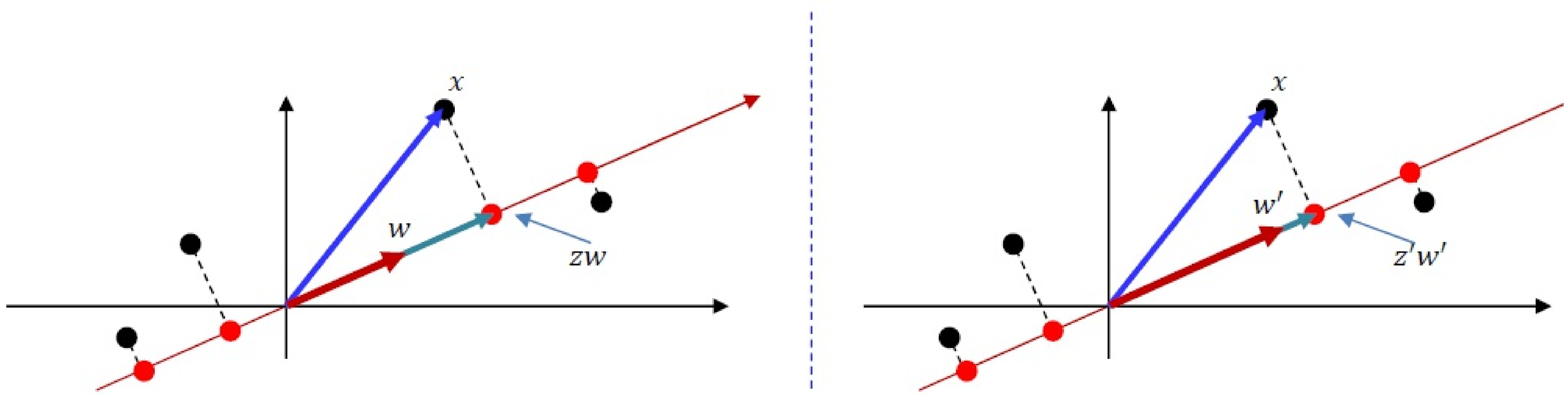
- PCA: Is actually a generative model for Gaussian data
  - Data lie close to a linear manifold, with orthogonal noise

Also Linear Transformation in Linear Algebra

VAE Intuition: PCA aka. Instance of LGM

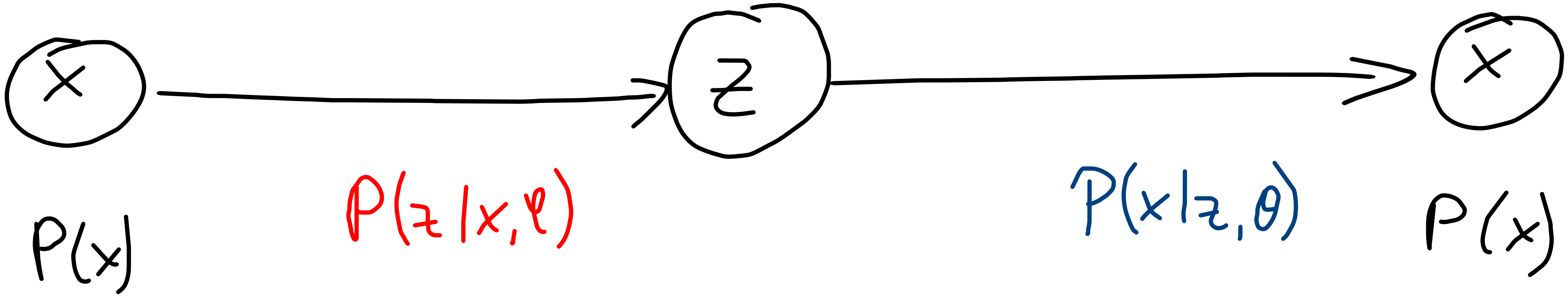


A minor issue: Scaling invariance

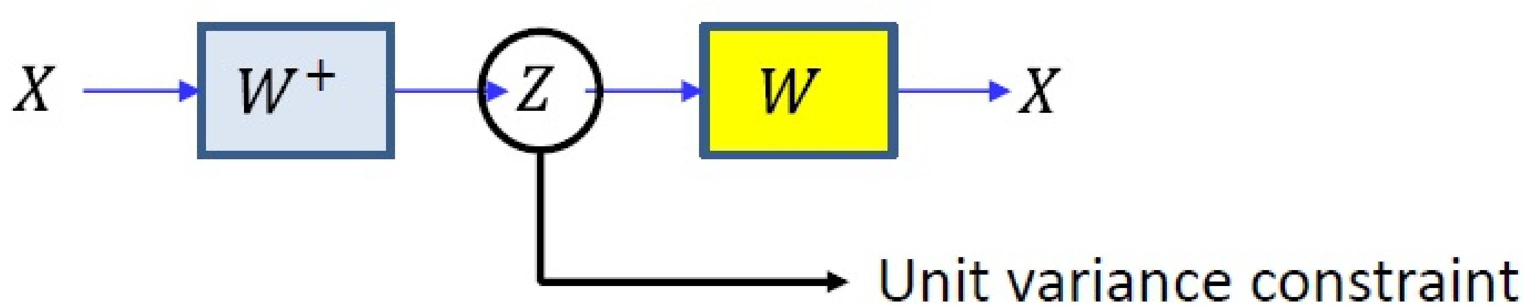


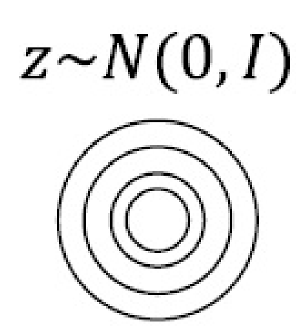
$$X \to \boxed{W^+} \to Z \to \boxed{W} \to X$$

VAE Intuition: PCA aka. Instance of LGM

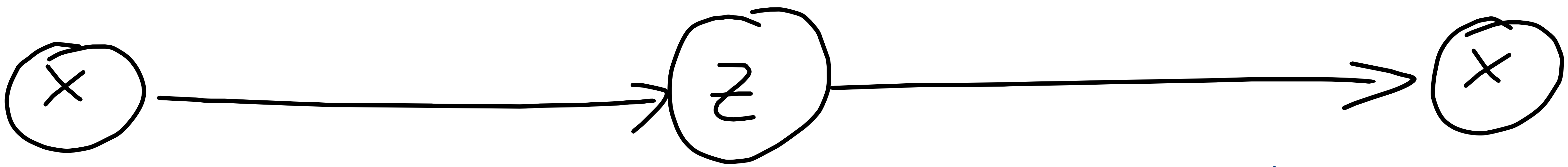


Constraining the linear AE



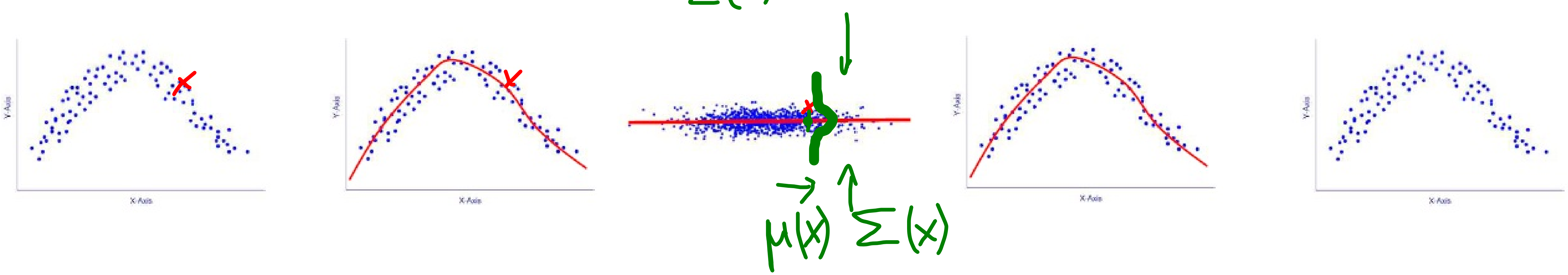


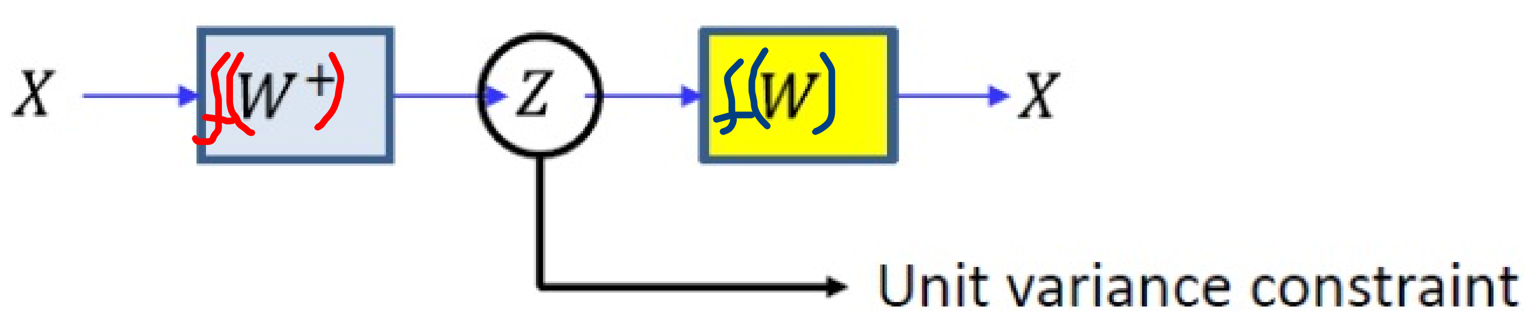
VAE Intuition: Non-Linear Gaussian Model



$P(x)$ $\quad P(z|x,\varphi) \quad$ $\mu(x) = E[z|x]$ $\quad P(x|z,\theta) \quad P(x)$

$\Sigma(x) = V(z|x)$



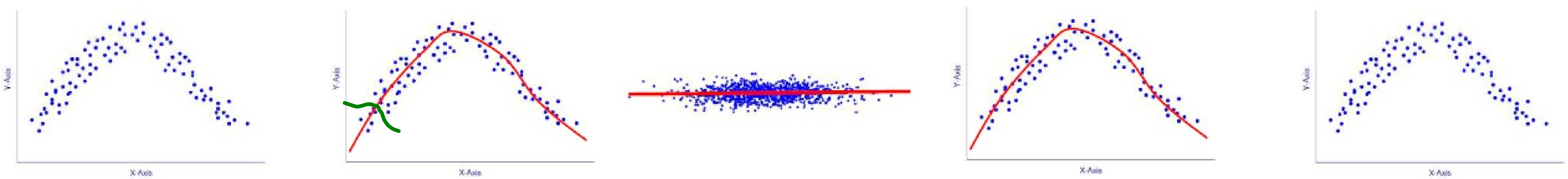


$f(\ldots) \rightarrow$ analytically intractable

VAE Intuition: Non-Linear Gaussian Model

$$x \longrightarrow z \longrightarrow \bar{x}$$

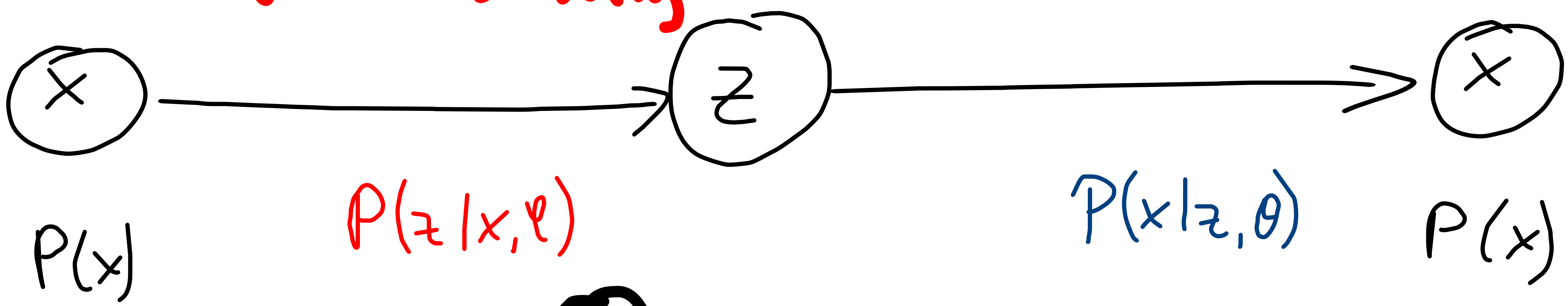
$P(x)$ $\quad$ $P(z|x,\varphi)$ $\quad$ $P(x|z,\theta)$ $\quad$ $P(x)$



Training a VAE: Step 1, 2 and 3

① Initialize Encoder Randomly

③ Train Decoder



② Sample z

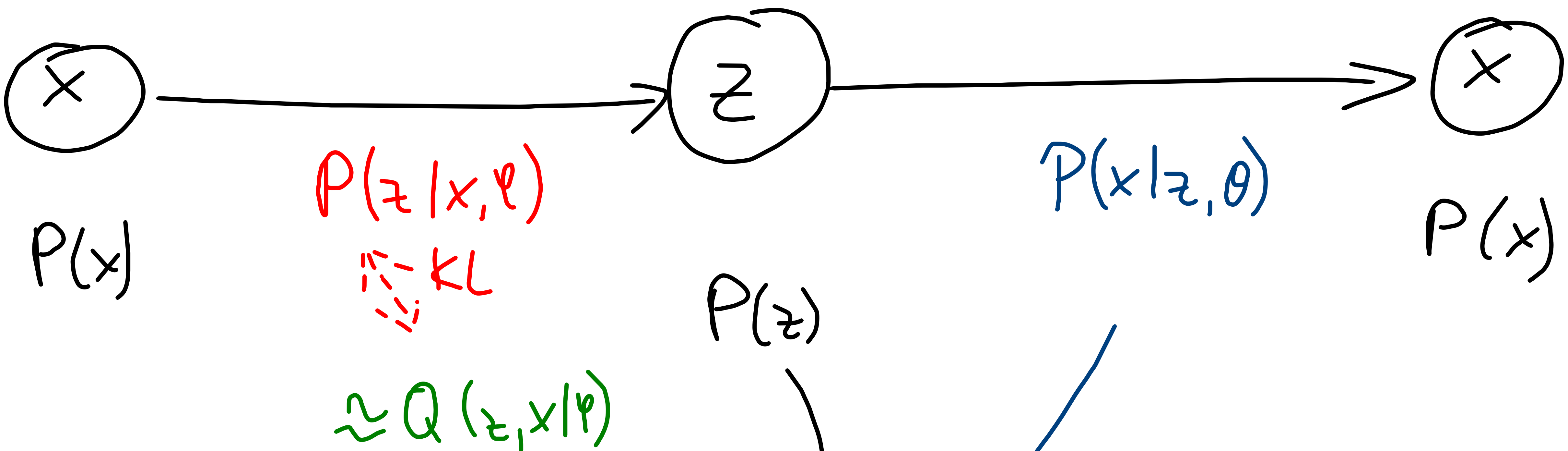
$$L(\theta,\sigma^2) = \sum_{(x,z)} \log P(x|z)$$

$$L(\theta,\sigma^2) = d\log\sigma^2 + \boxed{\frac{1}{\sigma^2}\sum_{(x,z)} \|x - f(z;\theta)\|^2}$$

$f(z)$: NN

Training a VAE: Step 4

4 Train Encoder



$$\varphi^* = \underset{\varphi}{\operatorname{argmin}}\, KL(Q(z,x), P(z|x))$$

$$\log Q(Z,X;\varphi) - \log P(Z|X;\theta)$$

$$= \sum_{(x,z)\in[X,Z]} \log Q(z,x;\varphi) - \underline{\log P(z|x;\theta)}$$

Bayes

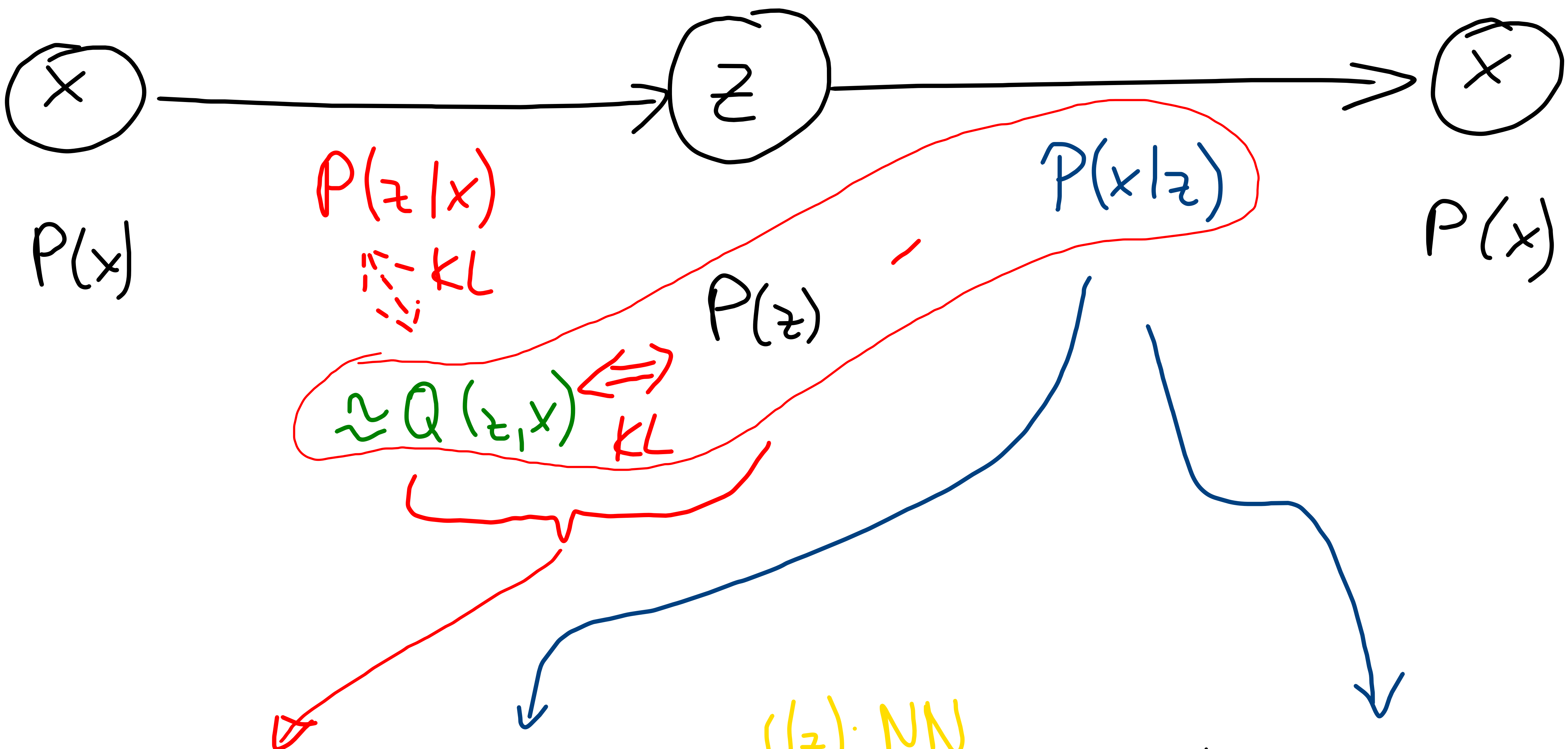
NLL

$$-\log P(z|x,\varphi) = -\log \frac{P(x|z,\theta)P(z)}{P(x)}$$

$$-\log P(z|x,\varphi) = -\log P(x|z,\theta) - \log P(z) + \cancel{\log P(x)}$$

Doesn't depend on φ or z

Training a VAE: Step 4



f(z): NN

$$\Rightarrow L_Q(\varphi) = \sum_{x\in X} KL(Q(z,x;\varphi),P(z)) - \boxed{\sum_{(x,z)\in[X,Z]} \log P(x|z;\theta)}$$

$$L(\theta,\sigma^2) = \sum_{(x,z)} \log P(x|z)$$

$$L_Q(\varphi) = \sum_{x\in X}\Big(tr\big(\Sigma(x;\varphi)\big) + \mu(x;\varphi)^T\big(\mu(x;\varphi) - d - \log|\Sigma(x;\varphi)|\big)\Big) \boxed{+ \frac{1}{\sigma^2}\sum_{(x,z)\in[X,Z]} \|(x - f(z;\theta))\|^2}$$

$$L(\theta,\sigma^2) = d\log\sigma^2 \boxed{+ \frac{1}{\sigma^2}\sum_{(x,z)} \|x - f(z;\theta)\|^2}$$