# Mean field

# Mean field

1. Select a family of distributions $Q$

$$Q = \{q \mid q(z) = \prod_{i=1}^{d} q_i(z_i)\}$$

MFA requires that each distribution in Q is factorized
over its individual latent variable dimensions (d).

2. Find best approximation $q(z)$ of $p^*(z)$:

$$\mathcal{KL}\left[q(z) \parallel p^*(z)\right] \to \min_{q \in Q}$$

by using KL! of course.

# Example

this is what we mean by factorization

vv

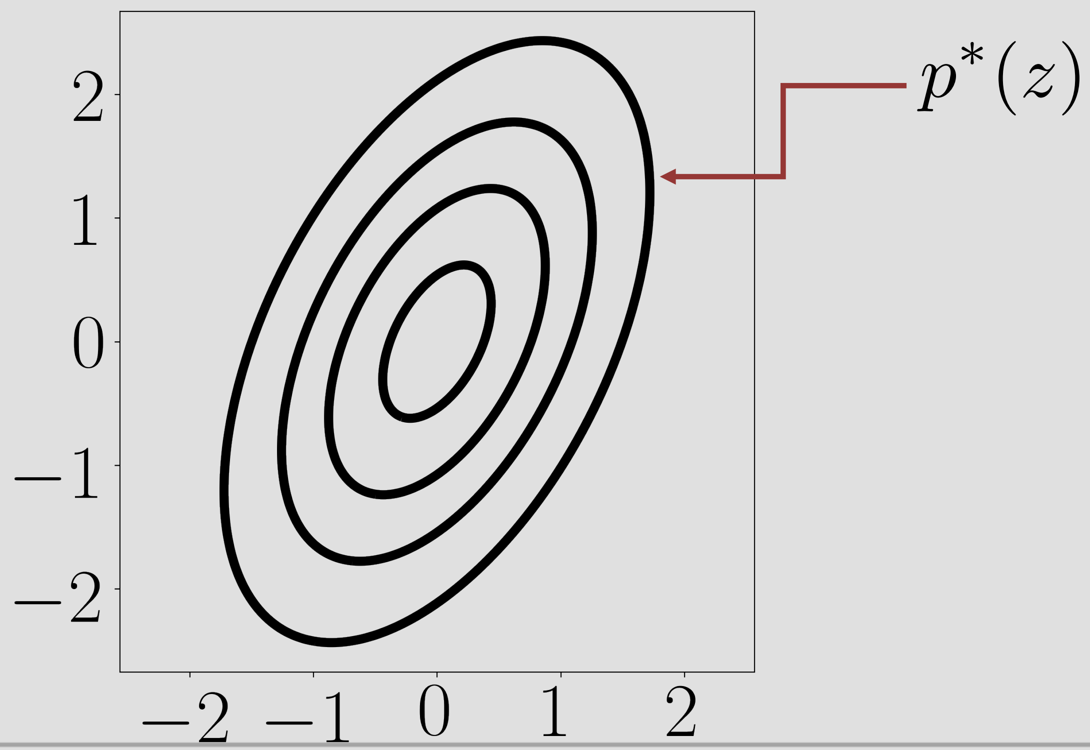
Suppose that our distribution's family is the normal distributions.

$$p^*(z_1, z_2) \approx q_1(z_1)q_2(z_2)$$

$$p^*(z_1, z_2) = \mathcal{N}(0, \Sigma)$$

Suppose that p*, our true posterior distribution, follows this normal:

$$q_1(z_1)q_2(z_2) = \mathcal{N}(0, \begin{pmatrix} \sigma_1^2 & 0 \\ 0 & \sigma_2^2 \end{pmatrix})$$



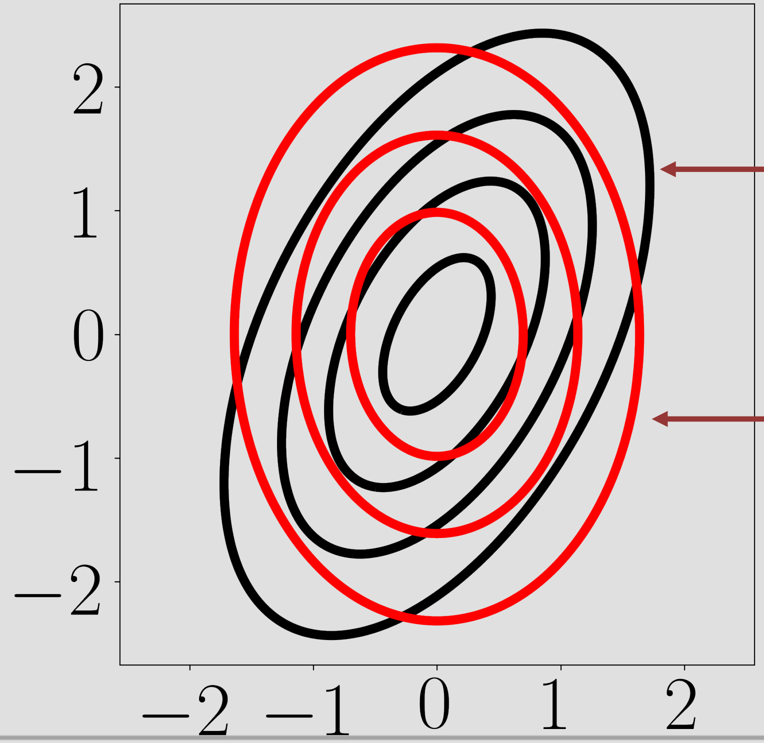
# Example

$$p^*(z_1, z_2) \approx q_1(z_1)q_2(z_2)$$

Then our approximated one would be like q(z) = q1(z1)q2(z2) (recall the factorization requirement)

$$p^*(z_1, z_2) = \mathcal{N}(0, \Sigma)$$

$$q_1(z_1)q_2(z_2) = \mathcal{N}(0, \begin{pmatrix} \sigma_1^2 & 0 \\ 0 & \sigma_2^2 \end{pmatrix})$$




$p^*(z)$

$q(z)$

# Optimization

$$\mathcal{KL}(q \,\|\, p^*) = \mathcal{KL}(\prod_{i=1}^{d} q_i \,\|\, p^*) \to \min_{q_1, q_2, \dots, q_d}$$

Coordinate descend: Do coordinate descent on individual factors, while keeping others constant.

1. $\mathcal{KL}(q \,\|\, p^*) \to \min_{q_1}$
2. $\mathcal{KL}(q \,\|\, p^*) \to \min_{q_2}$
3. ...

# Технический слайд (<= 12.5 min)

На доске вывод основной формулы + conditional conj.

$$\sum_X \prod_i q_i(x_i) \log \frac{\prod_i q_i(x_i)}{p(x)} = \sum_{x_j} \sum_{X_{-j}} q_j(x_j) \prod_{i\neq j} q_i(x_i) \left[\sum_k \log q_k(x_k) - \log p(x)\right] =$$

$$= \sum_{x_j} q_j(x_j) \sum_{X_{-j}} \prod_{i\neq j} q_i(x_i) \left[\sum_{k\neq j} \log q_k(x_k) + \log q_j(x_j)\right] - \sum_{x_j} q_j(x_j) \sum_{X_{-j}} \prod_{i\neq j} q_i(x_i) \log p(x) =$$

$$= \sum_{x_j} q_j(x_j) \cancel{\sum_{X_{-j}} \prod_{i\neq j} q_i(x_i)} \log q_j(x_j) - \sum_{x_j} q_j(x_j) \mathbb{E}_{X_{-j}} \log p(x) + const =$$

$$= \sum_{x_j} q_j(x_j) \left[\log q_j(x_j) - \mathbb{E}_{X_{-j}} \log p(x)\right] + const \longrightarrow \min_{q_j(x_j)}$$

$$\log q_j(x_j) = \mathbb{E}_{X_{-j}} \log p(x) + const.$$

$$q_j(x_j) = \frac{1}{Z} \exp\left(\mathbb{E}_{X_{-j}} \log p(x)\right)$$

TODO: the technical slide stuff.
NOT THIS ONE! These are shit notes.