Variational Inference

Variational inference

1. Select a family of distributions Q

Variational inference

1. Select a family of distributions Q

Example: $$\mathcal{N}\left(\mu, \begin{pmatrix} \sigma_1^2 & & & 0 \\ & \sigma_2^2 & & \\ 0 & & \ddots & \\ & & & \sigma_d^2 \end{pmatrix}\right)$$

Variational inference (ТЕХНИЧЕСКИЙ СЛАЙД)

1. Select a family of distributions Q

Example: $\mathcal{N}(\mu, \begin{pmatrix} \sigma_1^2 & & 0 \\ & \sigma_2^2 & \\ 0 & & \ddots & \\ & & & \sigma_d^2 \end{pmatrix})$

VARIATIONAL FAMILY

Variational inference

1. Select a family of distributions Q

Example: $\mathcal{N}(\mu, \begin{pmatrix} \sigma_1^2 & & & 0 \\ & \sigma_2^2 & & \\ & & \ddots & \\ 0 & & & \sigma_d^2 \end{pmatrix})$

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

Variational inference

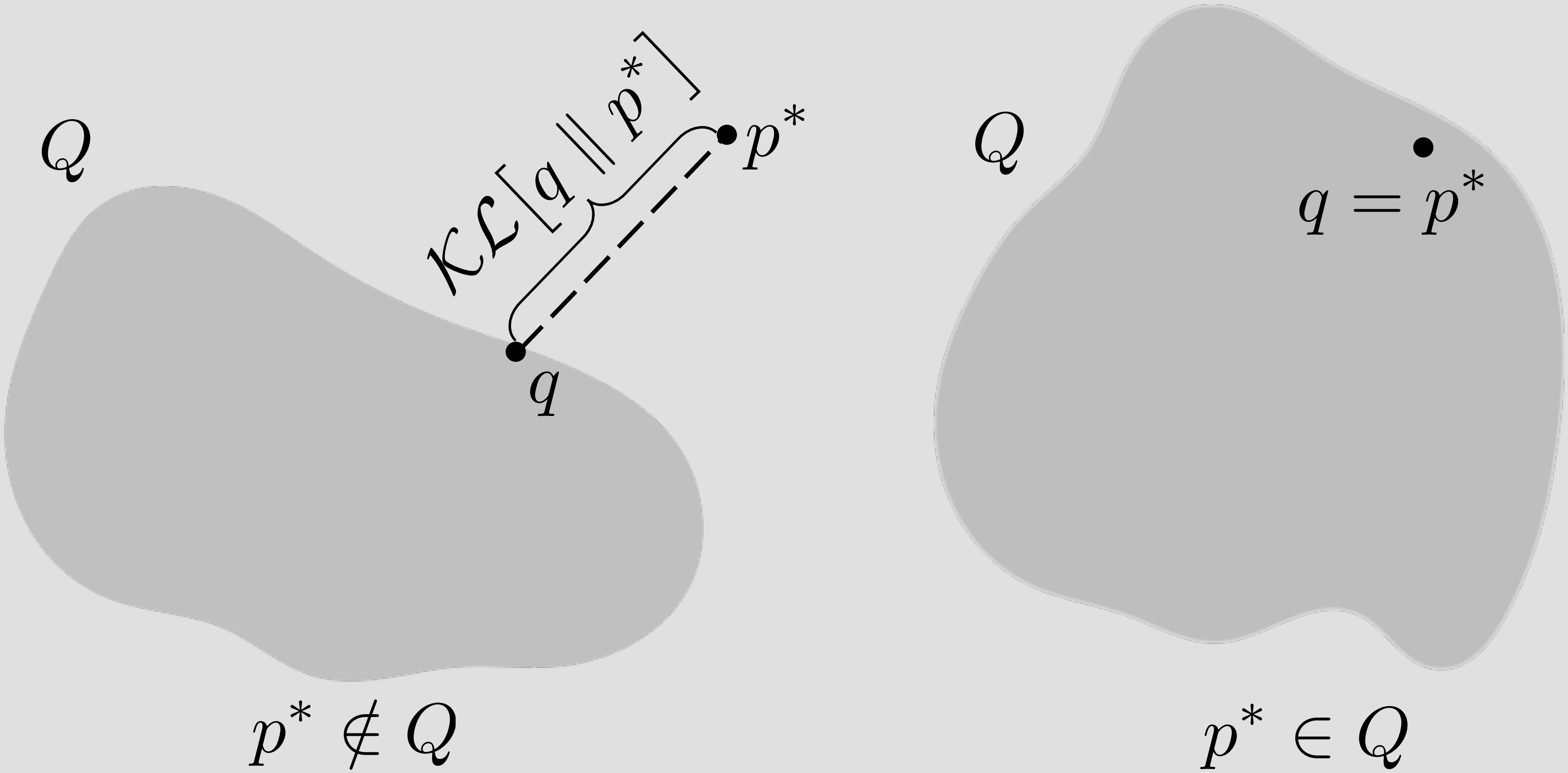
1. Select a family of distributions Q

Example: $\mathcal{N}\left(\mu, \begin{pmatrix} \sigma_1^2 & & & 0 \\ & \sigma_2^2 & & \\ & & \ddots & \\ 0 & & & \sigma_d^2 \end{pmatrix}\right)$

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

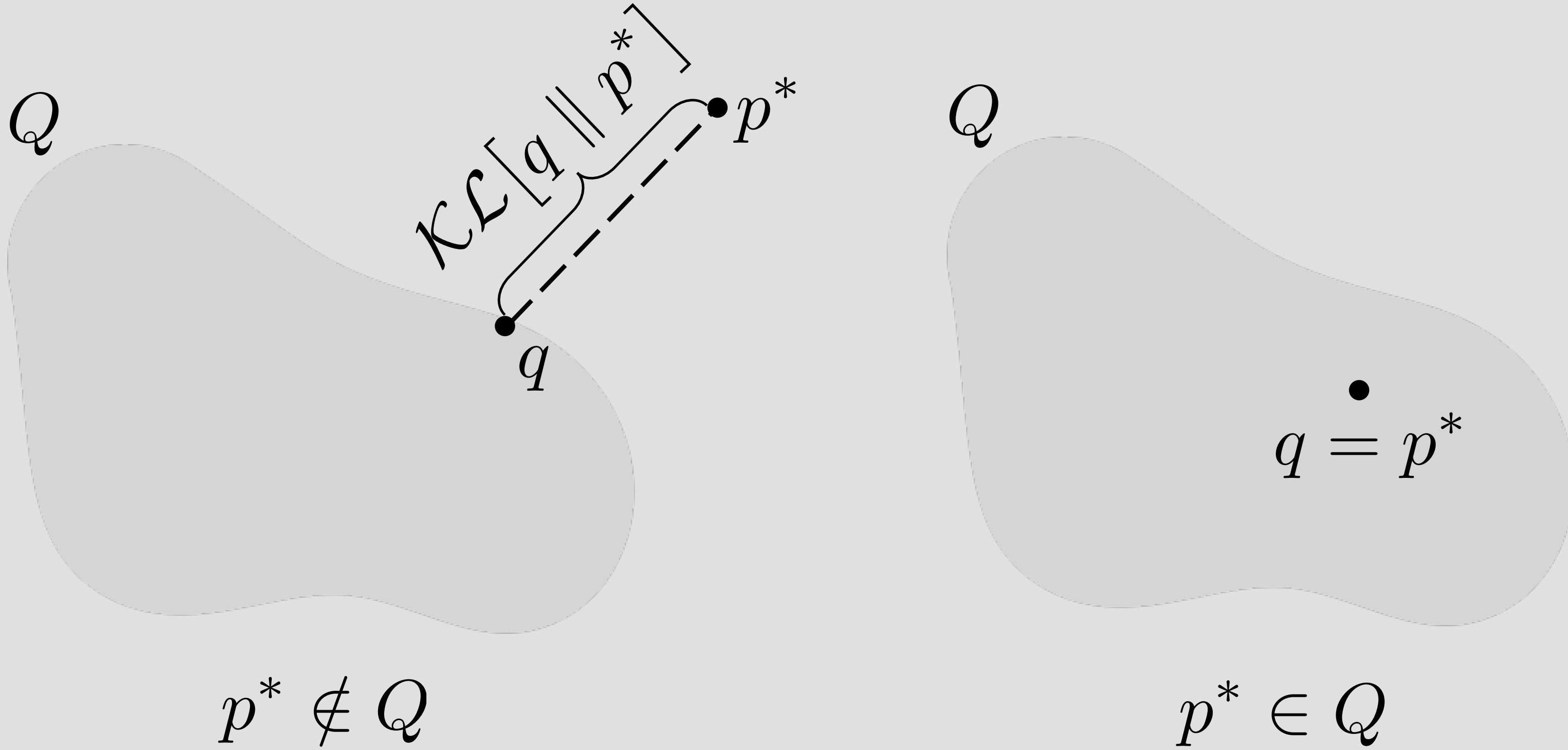
$$\mathcal{KL}\left[q(z) \parallel p^*(z)\right] \to \min_{q \in Q}$$

Choice of variational family



Choice of variational family (ТЕХНИЧЕСКИЙ СЛАЙД)

В стиле если Q не накроет p*, то будет расстояние какое-то.
А если накроет, то q и p* совпадут + немного темнее фигуру



LARGER Q => HARDER

Unnormalized distribution

$$p^*(z) = p(z|X) = \frac{p(X|z)p(z)}{p(X)} = \frac{\widehat{p}(z)}{Z}$$

Unnormalized distribution

$$p^*(z) = p(z|X) = \frac{p(X|z)p(z)}{p(X)} = \frac{\widehat{p}(z)}{Z}$$

Optimization:

Unnormalized distribution

$$p^*(z) = p(z|X) = \frac{p(X|z)p(z)}{p(X)} = \frac{\widehat{p}(z)}{Z}$$

Optimization:

$$\mathcal{KL}\left[q(z) \,\|\, \frac{\widehat{p}(z)}{Z}\right] =$$

Unnormalized distribution

$$p^*(z) = p(z|X) = \frac{p(X|z)p(z)}{p(X)} = \frac{\widehat{p}(z)}{Z}$$

Optimization:

$$\mathcal{KL}\left[q(z) \;||\; \frac{\widehat{p}(z)}{Z}\right] = \int q(z) \log \frac{q(z)}{\widehat{p}(z)/Z} dz$$

Unnormalized distribution

$$p^*(z) = p(z|X) = \frac{p(X|z)p(z)}{p(X)} = \frac{\widehat{p}(z)}{Z}$$

Optimization:

$$\mathcal{KL}\left[q(z) \parallel \frac{\widehat{p}(z)}{Z}\right] = \int q(z) \log \frac{q(z)}{\widehat{p}(z)/Z} dz$$

$$= \int q(z) \log \frac{q(z)}{\widehat{p}(z)} dz + \int q(z) \log Z dz$$

Unnormalized distribution

$$p^*(z) = p(z|X) = \frac{p(X|z)p(z)}{p(X)} = \frac{\widehat{p}(z)}{Z}$$

Optimization:

$$\mathcal{KL}\left[q(z) \parallel \frac{\widehat{p}(z)}{Z}\right] = \int q(z) \log \frac{q(z)}{\widehat{p}(z)/Z} dz$$

$$= \int q(z) \log \frac{q(z)}{\widehat{p}(z)} dz + \int q(z) \log Z dz$$

$$= \mathcal{KL}\left[q(z) \parallel \widehat{p}(z)\right] + \log Z$$

Unnormalized distribution

$$p^*(z) = p(z|X) = \frac{p(X|z)p(z)}{p(X)} = \frac{\widehat{p}(z)}{Z}$$

Optimization:

$$\mathcal{KL}\left[q(z) \parallel \frac{\widehat{p}(z)}{Z}\right] = \int q(z) \log \frac{q(z)}{\widehat{p}(z)/Z} dz$$

$$= \int q(z) \log \frac{q(z)}{\widehat{p}(z)} dz + \int q(z) \log Z dz$$

$$= \mathcal{KL}\left[q(z) \parallel \widehat{p}(z)\right] + \log Z$$

$$\mathcal{KL}\left[q(z) \parallel \widehat{p}(z)\right] \to \min_z$$