

Lecture 28:

Variational Inference

Professor Ilias Bilonis

Posterior approximations

The mean-field Gaussian approximation

$$q(z; \varphi) = \mathcal{N}\left(z \mid \mu, \text{diag}(\sigma_1^2, \dots, \sigma_k^2)\right)$$

$$\text{diag}(e^{\omega_1}, \dots, e^{\omega_k})$$

$$\sigma_i = \exp\{\omega_i\}$$

$$q(z; \varphi) = \mathcal{N}\left(z \mid \mu, \text{diag}(e^{\omega_1}, \dots, e^{\omega_k})\right)$$

$$\varphi = \{\mu, \omega_1, \dots, \omega_k\}$$

The full-rank Gaussian variational approximation

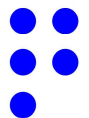
$$q(z; \varphi) = \mathcal{N}(z | \mu, \Sigma)$$

Cholesky of Σ

$$L = \begin{pmatrix} * & * & * \\ * & * & * \\ * & * & * \end{pmatrix}$$

(Note: In the original image, the upper triangular part of L is crossed out with a red line, and the lower triangular part is highlighted with a blue triangle.)

$$\varphi = \left\{ \mu, \underbrace{\text{red oval}}_{K \rightarrow}, \underbrace{\text{blue oval}}_{\frac{K^2 - K}{2}} \right\}$$



The low-rank Gaussian variational approximation

Full rank approx. not good for high dim. x .

$$q(z; \varphi) = N(z \mid \mu, \text{diag}(e^{\omega_1}, \dots, e^{\omega_K}) + U \cdot U^T)$$

$U \in \mathbb{R}^{K \times R}$
Rank

$$\varphi = \left\{ \underbrace{\mu}_K, \underbrace{\omega_1, \dots, \omega_K}_K, \underbrace{U}_{KR} \right\}$$