

Advanced probabilistic methods - Sketch

Christian Segercrantz 481056

March 24, 2022

$$p(\psi, \mathbf{Z}, \mathbf{W}, \mathbf{X}) = p(\mathbf{X} | \psi, \mathbf{Z}, \mathbf{W}) p(\psi) p(\mathbf{Z}) p(\mathbf{W}) \quad (1)$$

$$= \mathcal{N}_D(\mathbf{x}_n | \mathbf{W} \mathbf{z}_n, \text{diag}(\psi)^{-1}) \prod_{d=1}^D q^*(\mathbf{w}_d) \prod_{n=1}^N q^*(\mathbf{z}_n) \prod_{d=1}^D q^*(\psi_d) \quad (2)$$

$$= \mathcal{N}_D(\mathbf{x}_n | \mathbf{W} \mathbf{z}_n, \text{diag}(\psi)^{-1}) \mathcal{N}_K(\mathbf{w}_d | \mathbf{0}, \alpha \mathbf{I}) \mathcal{N}_K(\mathbf{z}_n | \mathbf{0}, \mathbf{I}) \text{Gamma}(\psi_d | a, b) \quad (3)$$

$$\mathbb{E}_{\psi, w} [\log q^*(\psi, \mathbf{Z}, \mathbf{W}, \mathbf{X})] = \mathbb{E}_{\psi, w} [\log (\mathcal{N}_D(\mathbf{x}_n | \mathbf{W} \mathbf{z}_n, \text{diag}(\psi)^{-1})) + \log (\mathcal{N}_K(\mathbf{w}_d | \mathbf{0}, \alpha \mathbf{I})) \quad (4)$$

$$+ \log (\mathcal{N}_K(\mathbf{z}_n | \mathbf{0}, \mathbf{I})) + \log (\text{Gamma}(\psi_d | a, b))] \quad (5)$$

$$\log (\mathcal{N}_D(\mathbf{x}_n | \mathbf{W} \mathbf{z}_n, \text{diag}(\psi)^{-1})) \propto -\frac{1}{2}(\mathbf{x}_n - \mathbf{W} \mathbf{z}_n)^\top \text{diag}(\psi)(\mathbf{x}_n - \mathbf{W} \mathbf{z}_n) \quad (6)$$

$$\log (\mathcal{N}_K(\mathbf{z}_n | \mathbf{0}, \mathbf{I})) \propto -\frac{1}{2} \mathbf{z}_n^\top \mathbf{z}_n \quad (7)$$

$$\mathbb{E}_{\psi, w} [\log q^*(\psi, \mathbf{Z}, \mathbf{W}, \mathbf{X})] \propto \mathbb{E}_{\psi, w} \left[-\frac{1}{2}(\mathbf{x}_n - \mathbf{W} \mathbf{z}_n)^\top \text{diag}(\psi)(\mathbf{x}_n - \mathbf{W} \mathbf{z}_n) - \frac{1}{2} \mathbf{z}_n^\top \mathbf{z}_n \right] \quad (8)$$

$$= \mathbb{E}_{\psi, w} \left[-\frac{1}{2}(\mathbf{z}_n^\top \mathbf{z}_n + \mathbf{x}_n^\top \text{diag}(\psi) \mathbf{x}_n - \mathbf{x}_n^\top \text{diag}(\psi) \mathbf{W} \mathbf{z}_n \right. \quad (9)$$

$$\left. - (\mathbf{W} \mathbf{z}_n)^\top \text{diag}(\psi) \mathbf{x}_n + (\mathbf{W} \mathbf{z}_n)^\top \text{diag}(\psi) \mathbf{W} \mathbf{z}_n) \right] \quad (10)$$

$$\propto \mathbb{E}_{\psi, w} \left[-\frac{1}{2}(\mathbf{z}_n^\top \mathbf{z}_n - 2 \mathbf{x}_n^\top \text{diag}(\psi) \mathbf{W} \mathbf{z}_n + \mathbf{z}_n^\top \mathbf{W}^\top \text{diag}(\psi) \mathbf{W} \mathbf{z}_n) \right] \quad (11)$$

$$= \mathbb{E}_{\psi, w} \left[-\frac{1}{2} \mathbf{z}_n^\top (\underbrace{\mathbf{I} + \mathbf{W}^\top \text{diag}(\psi) \mathbf{W}}_A) \mathbf{z}_n + \underbrace{\mathbf{x}_n^\top \text{diag}(\psi) \mathbf{W}}_{b^\top} \mathbf{z}_n \right], \quad \text{Completing the square} \quad (12)$$

$$= \mathbb{E}_{\psi, w} \left[\frac{1}{2}(\mathbf{z}_n - A^{-1}b)^\top A(\mathbf{z}_n - A^{-1}b) - \frac{1}{2} b^\top A^{-1}b \right] \quad (13)$$

$$= \frac{1}{2}(\mathbf{z}_n - \mathbb{E}_{\psi, w}[A^{-1}b])^\top \mathbb{E}_{\psi, w}[A](\mathbf{z}_n - \mathbb{E}_{\psi, w}[A^{-1}b]) - \frac{1}{2} \mathbb{E}_{\psi, w}[b^\top A^{-1}b] \quad (14)$$

$$\propto \frac{1}{2}(\mathbf{z}_n - \mathbb{E}_{\psi, w}[A^{-1}b])^\top \mathbb{E}_{\psi, w}[A](\mathbf{z}_n - \mathbb{E}_{\psi, w}[A^{-1}b]) \quad (15)$$

$$A^{-1} = \left(\mathbf{I} + \underbrace{\mathbf{W}^\top}_{\mathbf{K} \times \mathbf{D}} \underbrace{\text{diag}(\psi)}_{\mathbf{D} \times \mathbf{D}} \underbrace{\mathbf{W}}_{\mathbf{D} \times \mathbf{K}} \right)^{-1} = \left(\mathbf{I} + \sum_{\mathbf{d}=1}^{\mathbf{D}} \psi_{\mathbf{d}} \mathbf{w}_{\mathbf{d}} \mathbf{w}_{\mathbf{d}}^\top \right)^{-1} \quad (16)$$

$$b = \mathbf{W}^\top \text{diag}(\psi) \mathbf{x}_n \quad (17)$$

$$\mathbb{E}_{\psi, \mathbf{w}}[\mu_n] = \mathbb{E}_{\psi, \mathbf{w}}[A^{-1}b] = \mathbb{E}_{\psi, \mathbf{w}}[A^{-1}] \mathbb{E}_{\psi, \mathbf{w}}[b] \quad (18)$$

$$\mathbb{E}_{\psi, \mathbf{w}}[A^{-1}] = \mathbb{E}_{\psi, \mathbf{w}} \left[\mathbf{I} + \sum_{\mathbf{d}=1}^{\mathbf{D}} \psi_{\mathbf{d}} \mathbf{w}_{\mathbf{d}} \mathbf{w}_{\mathbf{d}}^\top \right]^{-1} \quad (19)$$

$$= \left(\mathbf{I} + \sum_{\mathbf{d}=1}^{\mathbf{D}} \mathbb{E}_{\psi}[\psi_{\mathbf{d}}] \mathbb{E}_{\mathbf{w}}[\mathbf{w}_{\mathbf{d}} \mathbf{w}_{\mathbf{d}}^\top] \right)^{-1} = \left(\mathbf{I} + \sum_{\mathbf{d}=1}^{\mathbf{D}} \langle \psi_{\mathbf{d}} \rangle \langle \mathbf{w}_{\mathbf{d}} \mathbf{w}_{\mathbf{d}}^\top \rangle \right)^{-1} \quad (20)$$

$$\mathbb{E}_{\psi, \mathbf{w}}[b] = \mathbb{E}_{\psi, \mathbf{w}} [\mathbf{W}^\top \text{diag}(\psi) \mathbf{x}] \quad (21)$$

$$= \mathbb{E}_{\mathbf{w}}[\mathbf{W}^\top] \mathbb{E}_{\psi}[\text{diag}(\psi)] \mathbf{x} \quad (22)$$

$$= \langle \mathbf{W}^\top \rangle \text{diag} \langle \psi \rangle \mathbf{x}_2 \quad (23)$$

$$\mathbb{E}_{\psi, w} [\log q^*(\psi, \mathbf{Z}, \mathbf{W}, \mathbf{X})] \propto \frac{1}{2}(\mathbf{z}_n - \underbrace{\mathbb{E}_{\psi, w}[A^{-1}] \mathbb{E}_{\psi, w}[b]}_{\mu_n})^\top \underbrace{\mathbb{E}_{\psi, w}[A^{-1}]}_{\mathbf{K}_n} (\mathbf{z}_n - \underbrace{\mathbb{E}_{\psi, w}[A^{-1}] \mathbb{E}_{\psi, w}[b]}_{\mu_n}) \quad (24)$$

$$\mathbb{E}_{\psi, w} [\log q^*(\psi, \mathbf{Z}, \mathbf{W}, \mathbf{X})] \propto \mathcal{N}(\mu_n | \mathbf{K}_n) \quad (25)$$

$$\mathbb{E}_{\psi, \mathbf{z}} [\log q^*(\psi, \mathbf{Z}, \mathbf{W}, \mathbf{X})] = \mathbb{E}_{\psi, \mathbf{z}} [\log \left(\prod_{n=1}^N \mathcal{N}_D(\mathbf{x}_n \mid \mathbf{W}\mathbf{z}_n, \text{diag}(\psi)^{-1}) \right) + \log (\mathcal{N}_K(\mathbf{w}_d \mid \mathbf{0}, \alpha \mathbf{I}))] \quad (26)$$

$$+ \log (\mathcal{N}_K(\mathbf{z}_n \mid \mathbf{0}, \mathbf{I})) + \log (\text{Gamma}(\psi_d \mid a, b)) \quad (27)$$

$$\propto \mathbb{E}_{\psi, \mathbf{z}} [\sum_{n=1}^N \log (\mathcal{N}_D(\mathbf{x}_n \mid \mathbf{W}\mathbf{z}_n, \text{diag}(\psi)^{-1})) + \log (\mathcal{N}_K(\mathbf{w}_d \mid \mathbf{0}, \alpha \mathbf{I}))] \quad (28)$$

$$\log (\mathcal{N}_D(\mathbf{x}_n \mid \mathbf{W}\mathbf{z}_n, \text{diag}(\psi)^{-1})) \propto -\frac{1}{2}(\mathbf{x}_n - \mathbf{W}\mathbf{z}_n)^\top \text{diag}(\psi)(\mathbf{x}_n - \mathbf{W}\mathbf{z}_n) \quad (29)$$

$$\log (\mathcal{N}_K(\mathbf{z}_n \mid \mathbf{0}, \mathbf{I})) \propto -\frac{1}{2}\mathbf{w}_d^\top (\alpha \mathbf{I}) \mathbf{w}_d \quad (30)$$

$$\mathbb{E}_{\psi, \mathbf{z}} [\log q^*(\psi, \mathbf{Z}, \mathbf{W}, \mathbf{X})] \propto \mathbb{E}_{\psi, \mathbf{z}} \left[-\frac{1}{2}(\mathbf{x}_n - \mathbf{W}\mathbf{z}_n)^\top \text{diag}(\psi)(\mathbf{x}_n - \mathbf{W}\mathbf{z}_n) - \frac{1}{2}\mathbf{w}_d^\top (\alpha \mathbf{I}) \mathbf{w}_d \right] \quad (31)$$