1 Computing a partition function

$$p(x;w) = \frac{1}{Z(w)} \exp\left(x^\top w\right) \tag{1}$$

$$\int p(x;w)dx = 1 \quad (x \in \mathbb{R}) \qquad \text{or} \qquad \sum_x p(x;w) = 1 \quad (x \in \mathbb{Q}) \tag{2}$$

$$\sum_x p(x;w) = \sum_x \frac{1}{Z(w)} \exp\left(x^\top w\right) \tag{3}$$

$$1 = \frac{1}{Z(w)} \sum_x \exp\left(x^\top w\right) \tag{4}$$

$$\log 1 = \log \frac{1}{Z(w)} \sum_x \exp\left(x^\top w\right) \tag{5}$$

$$0 = -\log Z(w) + \log \sum_x \exp\left(x^\top w\right) \tag{6}$$

$$\boxed{\log Z(w) = \log \sum_x \exp\left(x^\top w\right)} \tag{7}$$

2 Log-Sum-Exp Trick

Given $x = \{x_1, x_2, ..., x_n\}$, where $n \in \mathbb{N}$

$$y = \ln \sum_{n=1}^{N} e^{x_n} \qquad \Leftrightarrow \qquad e^y = \sum_{n=1}^{N} e^{x_n} \tag{8}$$

$$e^{-a} e^y = e^{-a} \sum_{n=1}^{N} e^{x_n} \qquad \Leftrightarrow \qquad e^{y-a} = \sum_{n=1}^{N} e^{x_n - a} \tag{9}$$

$$y - a = \ln \sum_{n=1}^{N} e^{x_n - a} \qquad \Leftrightarrow \qquad y = a + \ln \sum_{n=1}^{N} e^{x_n - a} \tag{10}$$

$$\boxed{\ln \sum_{n=1}^{N} e^{x_n} = a + \ln \sum_{n=1}^{N} e^{x_n - a}, \qquad \text{where} \qquad a = \max(x_i), \quad i = \overline{1..n}} \tag{11}$$

3 Compute a log of Gaussian PDF

Given $\mathcal{D} = (x^{(1)}, x^{(2)}, ...x^{(N)}), \quad \mathcal{D} \in \mathbb{R}^{(N,d)}$

$$\begin{aligned}\ln p(\mathcal{D}; (m, S)) = \ln \prod_{i=1}^{N} p(x^{(i)}; (m, S)) = \sum_{i=1}^{N} \ln p(x^{(i)}; (m, S)) = \\ -\frac{1}{2} \sum_{i=1}^{N} d \ln 2\pi + \ln |S| + (x^{(i)} - m)^\top S^{-1} (x^{(i)} - m) = \\ -\frac{1}{2} \Big(N d \ln 2\pi + N \ln |S| + \sum_{i=1}^{N} (x^{(i)} - m)^\top S^{-1} (x^{(i)} - m) \Big)\end{aligned} \tag{12}$$

$$\overline{\mathcal{D}} = \mathcal{D} - m[\text{np.newaxis}], \quad \overline{\mathcal{D}} \in \mathbb{R}^{(N,d)} \tag{13}$$

$$x^{(i)} = \mathrm{tr}\Big\{ x^{(i)} \Big\} \tag{14}$$

$$\mathrm{tr}\Big\{ ABC \Big\} = \mathrm{tr}\Big\{ BCA \Big\} \tag{15}$$

$$\begin{aligned}\sum_{i=1}^{N} (x^{(i)} - m)^\top S^{-1} (x^{(i)} - m) = \sum_{i=1}^{N} \mathrm{tr}\Big\{ (x^{(i)} - m)^\top S^{-1} (x^{(i)} - m) \Big\} = \\ \mathrm{tr}\Big\{ S^{-1} \sum_{i=1}^{N} (x^{(i)} - m)(x^{(i)} - m)^\top \Big\} = \mathrm{tr}\Big\{ S^{-1} \overline{\mathcal{D}}^\top \overline{\mathcal{D}} \Big\} = \mathrm{tr}\Big\{ \overline{\mathcal{D}} S^{-1} \overline{\mathcal{D}}^\top \Big\}\end{aligned} \tag{16}$$

$$SA = \overline{\mathcal{D}}^\top \quad \Leftrightarrow A = S^{-1} \overline{\mathcal{D}}^\top \tag{17}$$

$$A = \text{np.linalg.solve}(S, \overline{\mathcal{D}}^\top) \tag{18}$$

$$\mathrm{tr}\Big\{ \overline{\mathcal{D}} S^{-1} \overline{\mathcal{D}}^\top \Big\} = \mathrm{tr}\Big\{ \overline{\mathcal{D}} A \Big\} \tag{19}$$