統計的機械学習
第九回　レポート ID: 01

37-196360　森田涼介

2019 年 6 月 12 日

あるデータ x_i について，その潜在変数を z_i とし，また，事前分布のパラメータを θ と表す。Jensen の不等式を用いることで，次のように変分下限 L が求まる。

$$
\begin{aligned}
\log p_\theta(x_i) &= \log \int p_\theta(x_i,\, z_i)\,\mathrm{d}z \qquad (\text{潜在変数 } z_i \text{ は観測できないため周辺化}) && (1)\\
&= \log \int q(z_i) \frac{p_\theta(x_i,\, z_i)}{q(z_i)}\,\mathrm{d}z && (2)\\
&= \log E_{q(z_i)}\left[\frac{p_\theta(x_i,\, z_i)}{q(z_i)}\right] && (3)\\
&\geq E_{q(z_i)}\left[\log \frac{p_\theta(x_i,\, z_i)}{q(z_i)}\right] \qquad (\because \text{Jensen の不等式}) && (4)\\
&= \int q(z_i) \log \frac{p_\theta(x_i,\, z_i)}{q(z_i)}\,\mathrm{d}z_i && (5)\\
&= \int q(z_i) \log \frac{p_\theta(x_i|z_i)p(z_i)}{q(z_i)}\,\mathrm{d}z_i && (6)\\
&\equiv L[q(z_i),\, \theta;\, x_i] && (7)
\end{aligned}
$$

これを用いると，複数のデータ $x_{1:n}$ があるとき，その生成確率の対数尤度と変分下限は次のようになる。

$$
\begin{aligned}
\log p_\theta(x_{1:n}) &= \log \prod_{i=1}^{n} p_\theta(x_i) && (8)\\
&= \sum_{i=1}^{n} \log p_\theta(x_i) && (9)\\
&\geq \sum_{i=1}^{n} L[q(z_i),\, \theta;\, x_i] && (10)
\end{aligned}
$$