Variational Bayes

- Analysis goals: point estimates + uncertainties
- Interpretable, can incorporate expert information
- Challenge: fast Bayesian inference
- Modern problems: often large data, large dimension

一个问题用bayes model来求解是比较好建模的，建模很灵活，处理起来很简洁，适合做硬统计的工作。但是现在的数据量很大，数据维度很高，问题维度高，大多数的bayes model，如果用MCMC来求解的话，速度会非常慢。如果在这些问题上做bayes统计，有很大难度。

Bayesian Inference

Bayesian Inference

- **Data** $\boldsymbol{x} = \{x_1, x_2, \cdots, x_n\}$
- **Parameters** $\boldsymbol{z} = \{z_1, z_2, \cdots, z_m\}$
- Given **Prior** $p(\boldsymbol{z})$, we compute **Posterior distribution** $p(\boldsymbol{z}|\boldsymbol{x})$
- Using Bayes' theorem

$$\overbrace{p(\boldsymbol{z}|\boldsymbol{x})}^{\text{posterior}} = \frac{p(\boldsymbol{x}, \boldsymbol{z})}{p(\boldsymbol{x})} = \frac{\overbrace{p(\boldsymbol{x}|\boldsymbol{z})}^{\text{likelihood}}\overbrace{p(\boldsymbol{z})}^{\text{prior}}}{\underbrace{p(\boldsymbol{x})}_{\text{evidence}}}$$

1. Build a model: choose prior and likelihood
2. Compute the posterior using the observed data
3. Report a summary, e.g. posterior means and variances

如果现在你有一个后验分布，在高维贝叶斯模型中，你需要去计算它的means和variances是非常困难的，因为需要去做一个积分，对一个高维度的分布积分，数值上是非常难求解的。

- Posterior has no closed form. Directly computing $p(z|x)$ might **NOT** be feasible
- Markov Chain Monte Carlo (MCMC)
  - Accurate but slow
- Instead: an optimization approach – approximate posterior with q^*
  - Specify a family of nice distributions Q over the the parameter z
  - Find $q^*(z) \in Q$ to approximate $p(z|x)$

$$q^*(z) = \arg\min_{q(z)\in Q} L\big(q(z), p(z|x)\big)$$

- Variational Bayes (VB): L is KL divergence

$$L\big(q(z), p(z|x)\big) = \mathrm{KL}\big(q(z)||p(z|x)\big)$$

大多数情况下，后验分布是没呀解析解的，所以我们可以用MCMC去近似求解，MCMC虽然有理论上的保证，有理论上的收敛性，但是它的计算缓慢。现在有第二种思路，把贝叶斯模型的计算问题转换为一个优化问题，用一些方法去大大去加速它的计算。用一个变分近似分布去近似后验分布，去优化这个变分近似分布的参数，直到找到一组优秀的参数 z 得到一个逼近后验分布 p(z | x) 的变分近似分布 q(z)。在KL Divergence的衡量下可知 D_{KL}(q(z) || p(z | x))最小时，两个分布最接近。