When the updating process is tractable, (the last step that using observation to update prior, and conditional probability).

Approaches:

- Gibbs sampling
- Langevin Monte Carlo
- Sampling from a Laplace approximation
- Bootstrap.

Example. (Binary Feedback)

The travel time example again. We let the graph to represent a binomial bridge with M stages. Let θ_e be an independent gamma-distribution with $\mathbb{E}[\theta_e] = 1$, $\mathbb{E}[\theta_e^2] = 1.5$, and the observation

$$y_e \mid \theta = \begin{cases} 1 & \text{with prob } \dfrac{1}{1 + \exp(\sum_{e \in x_t} \theta - M)} \\ 0 & \text{o.w.} \end{cases}$$

$r_t = r(y_t) = y_t$. $\mathbb{E}\left[\sum_{e \in x_t} \theta\right] = M$.

---

* This model prevents us from using the conjugate properties because The gamma-distribution is not a conjugate prior of $y_e \mid \theta$.

* Bayesian Inference is also not easy. because when (x_t, y_t) are observed,

$$\begin{aligned} \hat{\theta}_{MAP} &= \arg\max_{\theta} P(\theta \mid x_t, y_t) \\ &= \arg\max_{\theta} P(y_t \mid x_t, \theta) \cdot \underbrace{P(\theta \mid x_t)}_{P(\theta)} \end{aligned}$$

= You need to manually take the derivative, and the result is complex.

Let f_{t-1} denote the posterior pdf of θ given a history of data $H_{t-1} = \{(x_i, y_i)\}_{i=1}^{t-1}$, i.e.

$$f_{t-1} = P(\theta \mid \mathcal{H}_{t-1})$$

Following TS framework.

1. Sampling $\hat{\theta}$ from f_{t-1}
2. Find the optimal x_t for deterministic $\hat{\theta}$, Apply x_t and observe y_t
3. Update to f_t

Here we introduce how to approximate the first step, i.e. sampling $\hat{\theta}$ from $\hat{f}_{t-1}$.

1. Gibbs Sampling.

Gibbs Sampling is a general Markov Chain Monte Carlo (MCMC) algorithm for drawing approximate samples from multivariate pdf.

Gibbs Sampling produces a sequence of sampled data

$\hat{\theta}^1, \hat{\theta}^2, \ldots$

forming a Markov Chain with a stationary distribution f_{t-1}. Under reasonable technical conditions (sufficient amount of samples), the limit distribution converges to f_{t-1}.

Gibbs Sampling. Framework to generate joint distribution $(x_1 .. x_n)$

1. Start with a random value in feasible region.
   $X_1 = (x_1, \cdots x_n)$
2. for round $r = 1, 2, \cdots$, do

   Given $x_2 \cdots x_n$, use $P_{x_1|x_2 \cdots x_n}$ to randomly choose x_1'

   Given $x_1, x_3, \cdots x_n$, use $P_{x_2|x_1, x_3 \cdots x_n}$ to randomly x_2'

   ⋮

   Given $x_1 \cdots x_{n-1}$, use $P_{x_n|x_1 \cdots x_{n-1}}$ to randomly choose a x_n'.

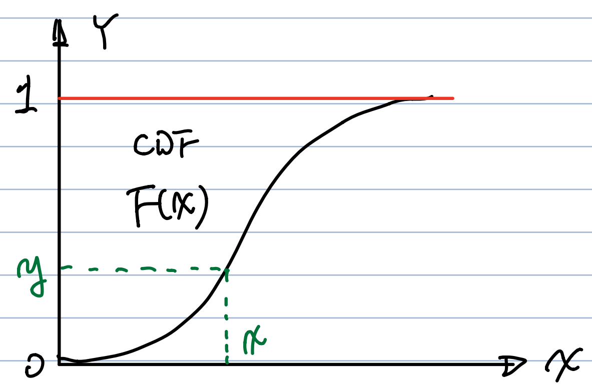
   Then our new sample this time is

   $$X_r = (x_1', x_2', \cdots, x_n')$$
3. When $r >$ some magic number.

The key point for using Gibbs Sampling is to get

$f = P_{x_r|(x_1 \cdots x_N)/x_r}$, and then get a sample $\sim f$

If we want an arbitrary sample with CDF F,



We sample y from a uniform distribution in $[0, 1]$, and then figure out the corresponding x from y. then $x \sim F$.

2. Laplace Approximation.

Let parameter $\theta \in R$ be a R.V. We have collected some data D. We want to get the posterior of θ

$$P(\theta | D) = \frac{1}{\Gamma} f(\theta)$$

Γ is the normalization term to make $\frac{1}{\Gamma} f(\theta)$ a valid pdf.
$f(\theta)$ is a function of θ

<u>Laplace Approximation framework</u>

① The second-order Taylor Expension.
Let θ_0 be the mode of of $P(\theta|D)$ where

$$\theta_0 = \arg\max_{\theta} P(\theta|D).$$

θ_0 can be found either analytically or numerically.
The let $g(\theta) = \log P(\theta|D)$, $g(\theta)$ around θ_0 can be approximated by.

$$g(\theta) \approx g(\theta_0) + g'(\theta_0)(\theta - \theta_0) + \frac{1}{2} g''(\theta_0)(\theta - \theta_0)^2$$

Since θ_0 is a maximum value. $g'(\theta_0) = \frac{d \log f(\theta)}{d\theta} = \frac{f'(\theta)}{f(\theta)} = \frac{0}{f(\theta)} = 0$

So $g(\theta) \approx g(\theta_0) + \frac{1}{2} g''(\theta_0)(\theta - \theta_0)^2$.

Then, we apply the exponential function to both side

$$\exp(g(\theta)) \approx \exp(g(\theta_0) + \frac{1}{2} g''(\theta_0)(\theta - \theta_0)^2)$$

$$\Rightarrow \quad \boxed{f(\theta) = f(\theta_0) \cdot \exp\left(\frac{1}{2} g''(\theta_0)(\theta-\theta_0)^2\right)} \quad \text{where } g(\theta) = \log f(\theta)$$

② The normalization constant.

$$\Gamma = \int_{\mathbb{R}} f(\theta)\, d\theta$$

$$= \int_{\mathbb{R}} f(\theta_0) \cdot \exp\left(\frac{1}{2} g''(\theta_0)(\theta-\theta_0)^2\right) d\theta$$

$$= f(\theta_0) \int_{\mathbb{R}} \exp\left(- \frac{(\theta-\theta_0)^2}{2 \cdot \left(-\frac{1}{g''(\theta_0)}\right)}\right) d\theta$$

looks like Normal Distribution ?

We make use of the Normal Distribution to compute this integration.

Let $\gamma = \sigma^{-2}$, then the pdf of $N(\mu, \sigma^2)$ satisfies.

$$1 = \int_{\mathbb{R}} \frac{1}{\sqrt{2\pi}\sigma} \exp\left(- \frac{(x-\mu)^2}{2\sigma^2}\right) dx$$

$$= \int_{\mathbb{R}} \frac{\sqrt{\gamma}}{\sqrt{2\pi}} \exp\left(- \frac{(x-\mu)^2}{2} \gamma\right) dx$$

$$\Rightarrow \int_{\mathbb{R}} \exp\left(- \frac{(x-\mu)^2}{2} \gamma\right) dx = \frac{\sqrt{2\pi}}{\sqrt{\gamma}}$$

Go back to our integration.

$$\Gamma = f(\theta_0) \cdot \int_{\mathbb{R}} \exp\left(- \frac{(\theta-\theta_0)^2}{2\left(-\frac{1}{g''(\theta_0)}\right)}\right) d\theta$$

$$\overset{\gamma = g''(\theta)}{=\!=} \quad f(\theta_0) \frac{\sqrt{2\pi}}{\sqrt{-g''(\theta)}}$$

③ Assembling All Parts

$$P(\theta|\mathcal{D}) = \frac{1}{\Gamma} f(\theta)$$

$$= \frac{\sqrt{-g''(\theta_0)}}{\cancel{f(\theta_0)}\sqrt{2\pi}} \cancel{f(\theta_0)} \exp\left(\frac{1}{2} g''(\theta)(\theta-\theta_0)^2\right)$$

$$= \frac{1}{\sqrt{2\pi}\sqrt{\frac{1}{-g''(\theta_0)}}} \exp\left(-\frac{(\theta-\theta_0)^2}{2\left(\frac{1}{-g''(\theta)}\right)}\right)$$

$$= \mathcal{N}\left(\theta_0, \frac{1}{-g''(\theta_0)}\right) \quad \text{where } g''(\theta_0) = \frac{d^2}{d\theta^2} \log f(\theta)\Big|_{\theta_0}$$

Then we can use a Gaussian to approximate the posterior.

---

3. Langevin Monte Carlo.

This is another MCMC but makes use of gradient information.

The idea is to sample the location of a partical doing Brownian motion in a restricted area, and the process is characterized by Langevin dynamics and define in differential equation.

Let $g(\theta)$ be the target pdf of θ (posterior we want), and we analyze its logarithm to make $\ln g(\theta)$ having more better properties (say L-smooth). So

$$g(\theta) = \frac{1}{\Gamma} e^{-\ln(g(\theta))}$$

and Γ is the normalization term

$$\Gamma = \int_\theta e^{-\ln(g(\theta))} d\theta$$

Denote $U(\theta) = -\ln(g(\theta))$ and usually θ is of very high dimension

We further assume $U(\theta)$ to be

① differentiable, i.e. $\nabla U(\theta)$ exists and can be efficiently computed.

② $U(\theta)$ is L-smooth: $\nabla^2 U(\theta)$ exists, and exists a sufficiently large L such that

$$\|\nabla U(\theta_1) - \nabla U(\theta_2)\| \leq L\|\theta_1 - \theta_2\| \text{ for any } \theta_1, \theta_2 \in \Theta$$

Langevin dynamics refer to the diffusion process.

$$d\theta_t = -\nabla U(\theta_t)\, dt + \sqrt{2}\cdot d \cdot B_t$$

standard Brownian Motion,

dimension of Brownian

Apply Euler-Maruyama to sample the diffusion path.

$$\phi_{n+1} = \phi_n + \left(\varepsilon \nabla \ln(g(\phi_n)) + \sqrt{2\varepsilon} W_n\right)$$

iid. stand Gaussian

Step size

$$\phi_{n+1} = \phi_n + \varepsilon A \nabla \ln(g(\phi_n)) + \sqrt{2}\sqrt{\varepsilon} A^{1/2} W_n$$

A is the PSD preconditioning matrix with

$$A = -\left(\nabla^2 \ln(g(\theta))\big|_{\theta=\theta_0}\right)^{-1}$$

negative inverse Hessian

where $\theta_0 = \arg\max_{\theta} \ln(g(\theta))$

4. Bootstrap Method.

Usually, bootstrap method is specific to a particular problem and usually not able to be generalized to more complex problem easily. Here is one example for Bernoulli Bandit Machine.

Like Laplace Approximation, we assume $\theta \in \mathbb{R}^d$. and we have historical data $H_{t-1} = \{(x_i, y_i)\}_{i=1}^{t-1}$, and $\hat{H}_{t-1}$ sampled uniformly with replacement from H_{t-1}.

key idea 1.

For Bernoulli model, the likelihood of θ given the historical data $\hat{H}_{t-1}$ for the shortest path recommendation problem (Binary feedback) described on the first page.

$$\hat{L}_{t-1}(\theta) = \prod_{\tau=1}^{t-1} \left(\frac{1}{1+\exp(\sum_{e \in \hat{x}_t} \theta_e - M)} \right)^{\hat{y}_t} \left(1 - \frac{1}{1+\exp(\sum_{e \in \hat{x}_t} e_t - M)}\right)^{1-\hat{y}_t}$$

We can use MLE to give an estimation of θ, but the problem to MLE is its relatively poor performance when t is small (not enough data), and MLE can not make use of prior info about $g(\theta)$ even if we have it.

The play around is as follows:

$$\hat{\theta} = \underset{\theta}{\operatorname{argmax}}\ e^{-(\theta-\theta^0)^T \Sigma (\theta - \theta^0)}\ \hat{L}_{t-1}(\theta)$$

Here θ^0 is a random sample from prior distribution f_0.

Σ is the covariance matrix of f_0.

This approximation utilizes the intuition that

$$\hat{L}_{t-1}(\theta) = Pr(\hat{H} \mid \theta).$$

to go from MLE to MAP, we need

$$\arg\max_{\theta} Pr(\hat{H} \mid \theta) \xrightarrow{P(\theta)} \arg\max_{\theta} Pr(\hat{H} \mid \theta) P(\theta)$$

$\frac{1}{\sqrt{(2\pi)^k |\Sigma|}} e^{-(\theta-\theta^0)^T \Sigma (\theta-\theta^0)}$ is a Gaussian with mean θ^0 and covariance Σ. Then $e^{-(\theta-\theta^0)^T \Sigma (\theta-\theta^0)} \hat{L}_{t-1}(\theta)$ is an approximation of posterior, which is also the approximation of prior for the next round $f_{t-1}(\theta)$. The reason we use Gaussian even through we know θ ~ Gamma is inspired by Laplace Approximation.

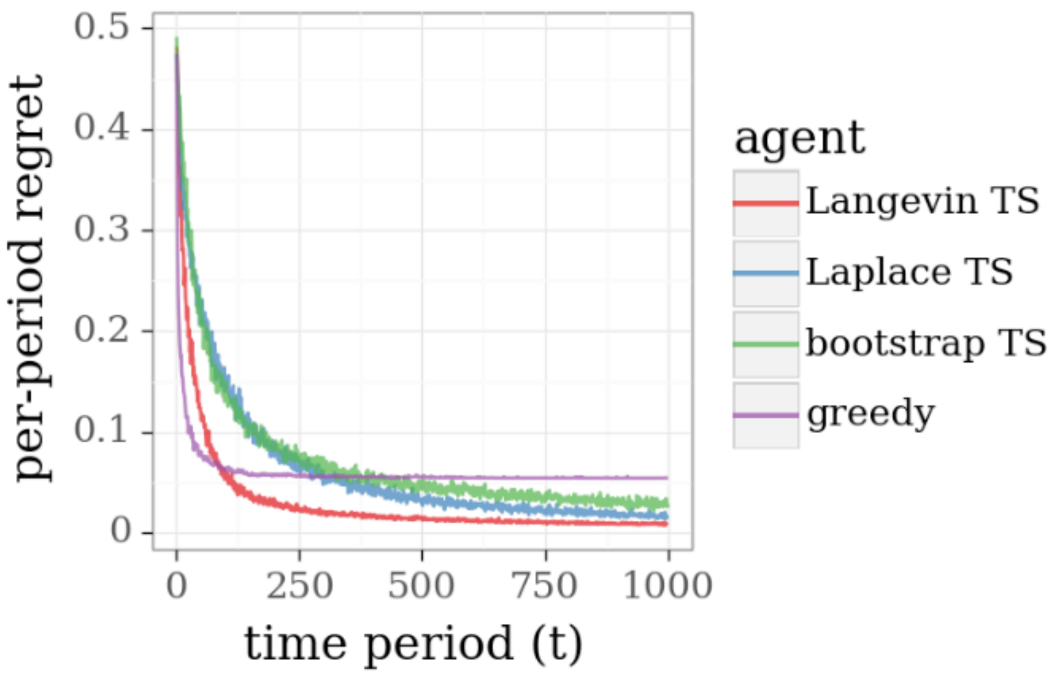
* When little data has been gathered, $\hat{\theta}$ is more like θ^0, and therefore scatered randomly among θ's support. This encourages the agent/player to explore more about the system.

* When more data has been gathered, the $\hat{\theta}$ is more determined by the likelihood term $\hat{L}_{t-1}(\theta)$, and the randomness most stems from the random selection of $\hat{H}_{t-1}$ from H_{t-1}

* In the shortest path problem, the approximated posterior, or $\hat{f}_{t-1}(\theta)$ is a log-concave function, therefore $\hat{\theta}$ can be efficiently computed using Newton' method with a backtracking line

search to maximize $\ln(\hat{f}_{t-1})$.

* For problems not easy to find the optimal $\hat{\theta}$, the framework can still be applied with local optimal or even an approximated maxima due to the nature of numerical iteration method.



The performance of four approximations for the binary feedback example.

From the performance, Bootstrapping works as well as Laplace.
The advantage of Bootstrapping is it's nonparametric, and work reasonably regardless of the form of posterior. It's nonparametric because it doesn't assume θ to be from Gaussian, but only take a random sample from a Gaussian, while Laplace always assume the posterior is Gaussian.
The disadvantage is that no guarantee of performance can be

achieved.