On Exponential Tilting

In Monte Carlo estimation, exponential tilting is a distribution shifting technique used in rare-event simulation, rejection and importance sampling.

1 Exponential Tilting

Given a random variable X with probability distribution P, density f, and m.g.f. $M_X(\theta) = E[e^{\theta X}] < \infty$, the exponentially titled measure P_θ is:

$$P_\theta(X \in dx) = \frac{E[e^{\theta X} 1[X \in dx]]}{M_X(\theta)} = e^{\theta x - \kappa\theta} P(X \in dx) \tag{1}$$

where $\kappa(\theta)$ is the c.g.f. $\log E[e^{\theta X}]$. Then P_θ has density

$$f_\theta(x) = \frac{e^{\theta x} f(x)}{M_X(\theta)} = \exp\{\theta x - \kappa(\theta)\} f(x). \tag{2}$$

2 Esscher Transform

Let $f(x)$ be a probability density. Then its Esscher transform is:

$$f_h(x) = \frac{e^{hx} f(x)}{\int_{-\infty}^{\infty} e^{hx} f(x) dx} \tag{3}$$

The effect of the Esscher transform on the normal distribution is a mean shifting:

$$E_h(\mathcal{N}(\mu, \sigma^2)) = \mathcal{N}(\mu + h\sigma^2, \sigma^2). \tag{4}$$

3 ReLU

Consider the distribution

$$p_\theta(X) = \frac{1}{Z(\theta)} \exp\left[f_\theta(X)\right] p_0(X) \tag{5}$$

where p_0 is Gaussian noise as reference distribution and $Z(\theta)$ denotes the partition function

$$Z(\theta) = \int \exp[f_\theta(X)] p_0(X) dX. \tag{6}$$

Then $p_\theta(Y)$ is an exponential tilting of p_0. In the trivial case of $f_\theta(X) = \theta X$, equation (5) is a Gibbs distribution and the effect of exponential tilting on $p_0(Y) \sim \mathcal{N}(\mu, \sigma^2)$ reduces to mean shifting $p_\theta(X) \sim \mathcal{N}(\mu + \theta\sigma^2, \sigma^2)$. If $f_\theta(X)$ is a ConvNet with piecewise linear rectification, then $f_\theta(X)$ is piecewise linear in X, and p_θ is piecewise mean shifted Gaussian.

References