GANs

Junrui Lin

2023 Mar

Method of Optimization
2023 Mar

Junrui Lin
Email: jl12680@nyu.edu

1 Setup

We have a structural model h associate with parameter θ, where we believe the data is generated through this model. Suppose we observe the sample

$$X_i = h(y_i, x_i)$$

And base on our model, we can generate our synthetic sample

$$X_i^\theta = h(y_i^\theta, x_i)$$

We have a family of discriminative models $D \in \mathcal{D}$

$$D : X_i \to [\varepsilon, 1 - \varepsilon]$$

Where $D(X_i) = P(X_i \text{ is drawn from sample})$, $1 - D(X_i) = P(X_i \text{ is synthetic})$
The estimator $\hat{\theta}$ is the solution of

$$\min_\theta \left\{ \max_{D \in \mathcal{D}} \frac{1}{N} \sum_{i=1}^{N} \log D(X_i) + \frac{1}{H} \sum_{j=1}^{H} \log[1 - D(X_j^\theta)] \right\}$$

2 Example

2.1 Logit Discriminator

The discriminator

$$D(X_i) = \log(\frac{1}{1 + e^{X_i'\beta}}) = P(X_i \text{ is drawn from sample})$$

$$\hat{\theta} = \arg\min_{\theta} \left\{ \max_{\beta} \frac{1}{N} \sum_{i=1}^{N} \log(\frac{1}{1+e^{X_i'\beta}}) + \frac{1}{H} \sum_{i=1}^{H} \log(\frac{e^{(X_i^{\theta})'\beta}}{1+e^{(X_i^{\theta})'\beta}}) \right\}$$

Suppose the data generating process is linear

$$y_i = x_i'\theta + \varepsilon_i,\ \ \varepsilon_i \sim N(0,1),\ \ X_i = (y_i, x_i y_i, x_i),\ \ X_i^{\theta} = (y_i^{\theta}, x_i y_i^{\theta}, x_i)$$

Step 0: Draw the random part of the model, in this case draw $\{\varepsilon_i^0\}_{i=1}^N$ i.i.d. from $N(0,1)$, here the simulation sample size is equal to the true sample size

Loop Start:

Step 1: First give inital guess θ^0, generate synthetic data by

$$y_i^{\theta^0} = x_i'\theta^0 + \varepsilon_i^0,\ \ X_i^{\theta^0} = (y_i^{\theta^0}, x_i y_i^{\theta^0}, x_i)$$

Step 2: Training the discriminator, get β^0 by

$$\beta^0 = \arg\max_{\beta} \left\{ \frac{1}{N} \sum_{i=1}^{N} \log(\frac{1}{1+e^{X_i'\beta}}) + \frac{1}{N} \sum_{i=1}^{N} \log(\frac{e^{(X_i^{\theta^0})'\beta}}{1+e^{(X_i^{\theta^0})'\beta}}) \right\}$$

Step 3: Updating the data generating parameter θ^1 by 1-step gradient descent

$$L(\theta^0, \beta^0) = \left\{ \frac{1}{N} \sum_{i=1}^{N} \log(\frac{1}{1+e^{X_i'\beta^0}}) + \frac{1}{N} \sum_{i=1}^{N} \log(\frac{e^{(X_i^{\theta^0})'\beta^0}}{1+e^{(X_i^{\theta^0})'\beta^0}}) \right\}$$

Notice that β is associate with θ, so we need to compute the derivative numerically. Little perturbation of θ won't affect β, because of Step 2 and Envelop Theorem.

We want to compute $\nabla_{\theta^0} L(\theta^0, \beta^0)$ using

$$\nabla_{\theta^0} L(\theta^0, \beta^0) = \frac{L(\theta^0 + h, \beta^0) - L(\theta^0 - h, \beta^0)}{2h}$$

Note: Here when compute $L(\theta^0 + h, \beta^0)$, $L(\theta^0 - h, \beta^0)$, we need to go back to Step 1 to get new X_i

Updating θ_1 by

$$\theta^1 = \theta^0 + \gamma \nabla_{\theta^0} L(\theta^0, \beta^0)$$

Where the step size γ is choosen by us, but γ is usually small.

Step 4: Iteration, go back to Step 1, $\theta^0 = \theta^1$...

Loop End: If the model is correctly specified, the parameters will converge. In this case $\beta^t \to 0$, $D(X) = 0.5$