Class 7

Shikhar Saxena

February 9, 2023

Contents

Comparison of ML and MAP 1

Generative modelling 1
- Estimate distribution by Counting 2
- Continuous modeling 2

Comparison of ML and MAP

- ML Estimate: max $f_{Y|X}(y|x)$
- MAP Estimate: max $f_{Y|X}(y|x)f_X(x)$

Generative modelling

Given obervations $x_1, x_2, \dots, x_n \sim p$. We want to sample from the unknown distribution p.

One possibility is ML Estimate (fitting a template distribution over the given data as best as it can). Once we have $\hat{p}$ we can sample from that.

Definition 1 (Generative Modeling)**.** *A procedure that produces $\hat{p}$ to approximate p that produces samples $\hat{x} \sim \hat{p}$.*

> **General Defintion:**
> Construct a function generator $g : Z \to X$ that maps source of simple randomness $Z \sim q$ to output $\hat{x} = g(z) \sim \hat{p}$ such that $\hat{p} \simeq p$.

Generative modeling is a *statistical problem* (because based on data).

Example 1. *Sample from Gaussian $p(x) = N(x; \mu, \sigma^2)$ given access to sample from $g(z) = N(z; 0, 1)$.*

$N(\mu, \sigma^2) \sim \sigma z + \mu \implies z \to \sigma z + \mu$

These kind of mappings are called pushforward distribution.

We want to sample from an easy distribution (primitive distribution) then there's a map that generates the cat image (or whatever distribution we want to generate).

Definition 2 (Pushforward Distribution)**.** *Given probability space* (Z, q) *a function* $g : Z \to X$ *induces a pushforward distribution on* X *defined for any set* $A \subset X$ *by*

$$P(A) = \int_{g^{-1}(A)} q(z)dz$$

Estimate distribution by Counting

Goal: Estimate p from samples $x^1, x^2, \ldots x^n \sim q$. Asume samples takes on values in a finite set $\{1, \ldots k\}$.

Define $p_i = P(i$ appears in the dataset$)$.

Example 2. *Roll die. Here* $k = 6$*.*

Let our samples be $\{1, 6, 2, 1, 5, 6, 3\}$*.*

So $p_1 = 2/7, p_2 = 1/7, p_3 = 1/7, p_4 = 0, p_5 = 1/7, p_6 = 2/7$*.*

This doesn't give us anything so far. We don't know how to generate anything right now. We want to go from $Z \to p$*. We already have* p*. Now we need a distribution* Z *and the function generator. So we can take* $U(0, 1)$ *or* $N(0, 1)$*.*

Use inverser transform $F^{-1}(U)$*. This will be our pushforward distribution to sample from* p*.*

Definition 3 (Inverse of CDF)**.**

$$g(z) = \inf\{x : F(x) \geq z\}$$

Given $Z \sim U(0, 1)$*, it follows that* $g(z) \sim p$*.*

Continuous modeling

$\hat{p}$ is defined as a pushforward of a density q by generator $g : Z \to X$

$$P(A) = P(g^{-1}(A)) = \int_{g^{-1}(A)} g(z)dz = \int_A q(g^{-1}(x)) \left|\nabla_x g^{-1}(x)\right| dx$$

This is difficult to compute. For simple generators g when the inverse and Jacobians are easy to compute, then $\hat{p}(x) = q(g^{-1}(x))|\nabla_x g^{-1}(x)|$ can be used to convert a generator to density estimator.

When the data is complicated then g will also mostly be complicated, so it will be difficult to use this approach.