

kernal density estimation on Riemannian manifolds

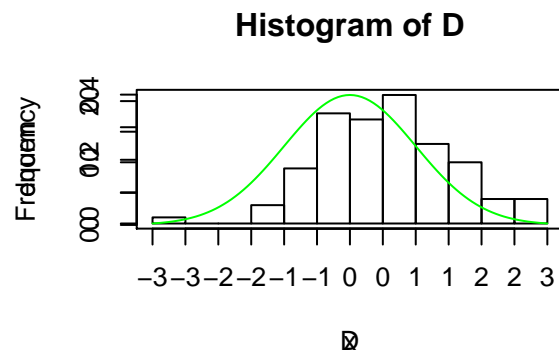
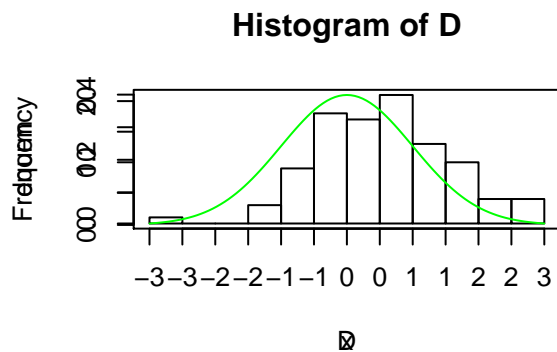
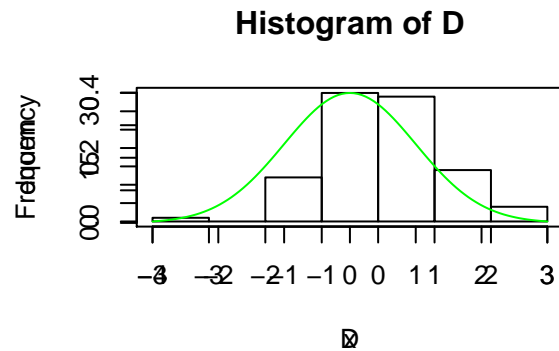
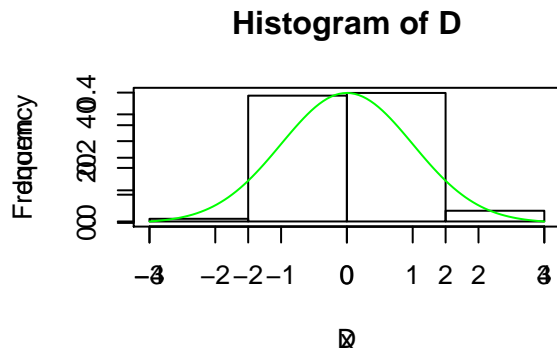
Wang Rui

November 29, 2016

density estimation

Naive estimator: histogram

```
D=rnorm(n = 100,mean = 0,sd = 1)
par(mfrow=c(2,2))
for(i in 1:4){
  hist(D,breaks=3*i)
  par(new=TRUE)
  plot(dnorm,xlim=c(-3,3),col="green")
}
```



- bandwidth is hard to choose
- break points are hard to choose
- effected by abnormal points
- **it is not smooth**

kernal estimation

For each sample point, we locate a small wavelet around it. The wavelet is known by kernal function in statistics.

Normal kernal

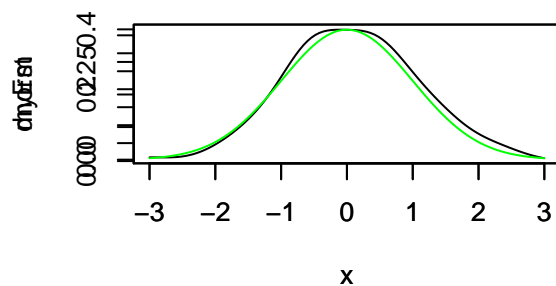
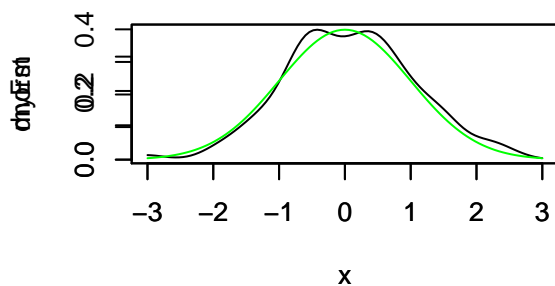
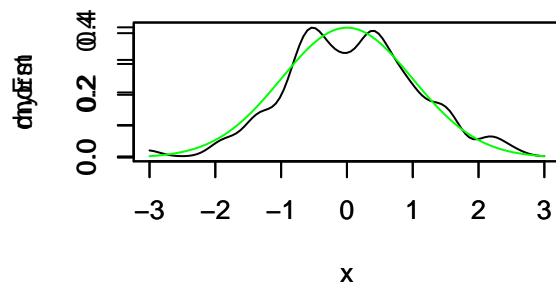
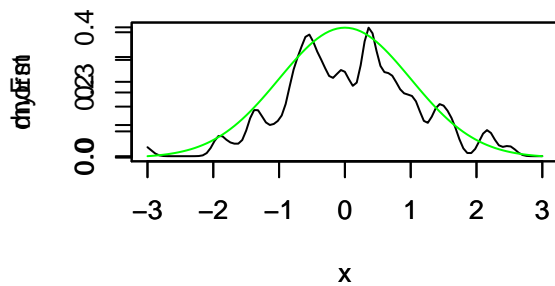
```

myF=list()
myBd=1
theIndex=1:length(D)
myF=lapply(theIndex,
            function(i) {
              function(x) dnorm(x,mean=D[i],sd=myBd)
            }
)

myEst=function(x){
  theS=0
  for(i in 1:length(myF)){
    theS=theS+myF[[i]](x)
  }
  theS/length(myF)
}

par(mfrow=c(2,2))
for(i in 1:4){
  myBd=0.1*i
  plot(myEst,xlim = c(-3,3))
  par(new=TRUE)
  plot(dnorm,xlim=c(-3,3),col="green")
}

```



```

#D=rnorm(n = 100,mean = 0,sd = 1)
myF=list()
theIndex=1:length(D)
bb=2
myF=lapply(theIndex,
            function(i) {
              function(x) dbeta(x-D[i]+0.5,bb,bb)
            }
)

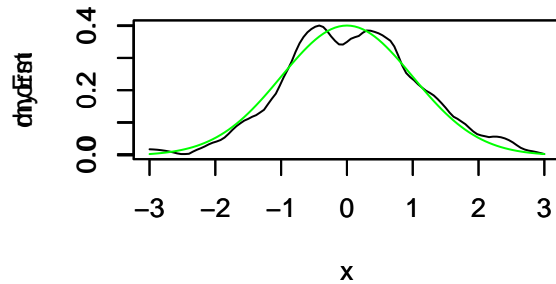
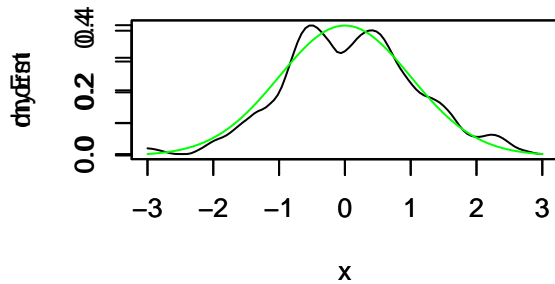
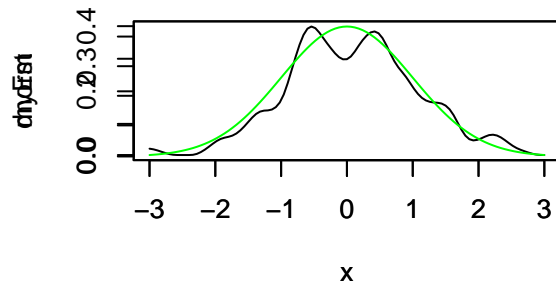
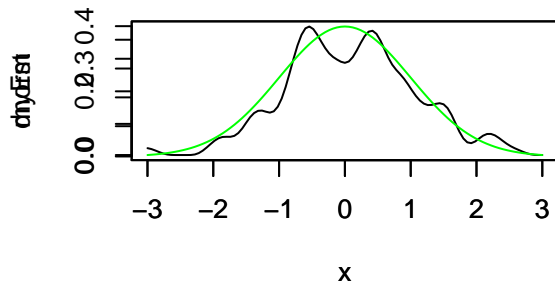
```

```

    }
)

myEst=function(x){
  theS=0
  for(i in 1:length(myF)){
    theS=theS+myF[[i]](x)
  }
  theS/length(myF)
}
par(mfrow=c(2,2))
for(bb in 5:2){
  plot(myEst,xlim = c(-3,3))
  par(new=TRUE)
  plot(dnorm,xlim=c(-3,3),col="green")
}

```



In R^d , define isotropic kernel by

1. $\int_{R^d} K(\|x\|) d\lambda(x) = 1$ (it's a probability density)
2. $\int_{R^d} x K(\|x\|) d\lambda(x) = 0$ (located at zero)
3. $\int_{R^d} \|x\|^2 K(\|x\|) d\lambda(x) < \infty$ (in L^2 space)
4. $\text{supp} K = [0, 1]$ (local defined)
5. $\sup K(x) = K(0)$

where $\lambda(x)$ is the Lebesgue measure.

In R^d , kernel density estimator is defined by

$$f_{n,K}(x) = \frac{1}{n} \sum_{i=1}^n \frac{1}{r^d} K\left(\frac{\|p - X_i\|}{r}\right)$$

The main task in the paper is to generalize things above from R^d to manifold.

exponential map

Let's turn to geometry. We assume the m -dimensional Riemannian Manifolds M is compact and complete. the true probability density on M is controlled by Riemannian measure. Then exponential map exists. $\exp_p(X)$ is a map from $T_p M$ to M . We choose the basis of $T_p M$ be the **normal** basis. Note that $T_p M$ is thus a Euclidean space. **So \exp_p^{-1} is actually a chart!** We will always use this chart.

We introduce some magical properties of the exponential chart!

The norm derivative of Geodesics $\gamma'(t)$ is a constant! So $\gamma'(t) = \gamma'(0)$. So the length of Geodesics is

$$\int_{[0,1]} \|\gamma'(t)\| dt = \|\gamma'(0)\|$$

So exponential map:

$$\gamma'(0) \rightarrow \gamma(1)$$

has the property:

$$\|\gamma'(0)\| = d(\gamma(1), p)$$

That is:

$$d_g(\exp_p(x), p) = \|x\|$$

It's very **important!**

Let μ_{g_p} be the Lebesgue measure on $T_p M$ (with normal chart). Define a measure on $T_p M$ by

$$\mu_{\exp_p^* g}(K) = \nu_g(\exp_p K)$$

So the Radon nikodym derivative

$$\frac{d\mu_{\exp_p^* g}}{d\mu_{g_p}}$$

is a function on $T_p(M)$. Define the **volume density function** θ_p by

$$\theta_p(q) = \frac{d\mu_{\exp_p^* g}}{d\mu_{g_p}}(\exp_p^{-1}(q))$$

which can be viewed as the change of variable from M to the chart $T_p M$.

The kernel density estimator is defined as

$$f_{n,K}(p) = \frac{1}{n} \sum_{i=1}^n \frac{1}{r^d} \frac{1}{\theta_{X_i}(p)} K\left(\frac{d_g(p, X_i)}{r}\right)$$

where r is the bandwidth.

To show it's indeed a probability density with respect to Riemannian measure. We need to integrate p and see if the result is equal to 1. Notice that the n terms are symmetry. We only need to compute

$$\int_M \frac{1}{r^d} \frac{1}{\theta_{X_1}(p)} K\left(\frac{d_g(p, X_1)}{r}\right) d\nu_g(p)$$

If we assume r is sufficiently small to contained in the local chart, we have

$$\int_M \frac{1}{r^d} \frac{1}{\theta_{X_1}(p)} K\left(\frac{d_g(p, X_1)}{r}\right) d\nu_g(p) = \int_{B_M(X_1, r)} \frac{1}{r^d} \frac{1}{\theta_{X_1}(p)} K\left(\frac{d_g(p, X_1)}{r}\right) d\nu_g(p)$$

It can be pull back to $T_p(M)$:

Let $\exp_{X_1}(x) = p$, then $d_g(p, X_1) = \|x\|$, and $B(0, r) \rightarrow B_M(X_1, r)$, and

$$\theta_{X_1}(p) = \frac{d\mu_{\exp_p^* g}}{d\mu_{g_p}}(x)$$

and

$$d\nu_g(p) = d\mu_{\exp_p^* g}(x)$$

So

$$\int_{B_M(X_1, r)} \frac{1}{r^d} \frac{1}{\theta_{X_1}(p)} K\left(\frac{d_g(p, X_1)}{r}\right) d\nu_g(p) = \int_{B(0, r)} \frac{1}{r^d} \frac{d\mu_{g_p}}{d\mu_{\exp_p^* g}}(x) K\left(\frac{\|x\|}{r}\right) d\mu_{\exp_p^* g}(x)$$

and is

$$\int_{B(0, r)} \frac{1}{r^d} K\left(\frac{\|x\|}{r}\right) d\mu_{g_p}(x) = 1$$

Consistency (Statistical property)

I will proof it on the blackboard if time promise.

Conclusion

The kernal density method can be generalized to manifold. It can be used in directional statistics as far as I know. I think the results in this paper can be generalized easily. For example, the L^2 convergence can be relaxed to uniformly convergence by the standard tool of Empirical process. As far as I know, such method is still in theory period, and have few application.

There's many more interesting problem in statistics such as estimating the manifold itself! Such direction is much more difficult than this paper. So I choose to report this paper not for it's an important problem (it's actually not very important in statistics), but just for it's a short paper and is relatively easy to follow!