Diffusion Kernels on q-Gaussian Manifold

Juan Carlos Arango Parra

May 11, 2018

Advisers: PhD Gabriel Ignacio Loaiza Ossa
PhD Carlos Alberto Cadavid Moreno

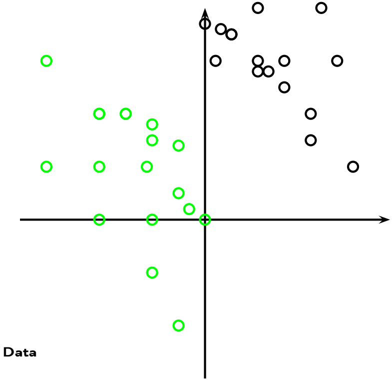
Doctoral Seminar 2
Universidad EAFIT
Department of Mathematical Sciences
PhD in Mathematical Engineering

Objetive of the presentation

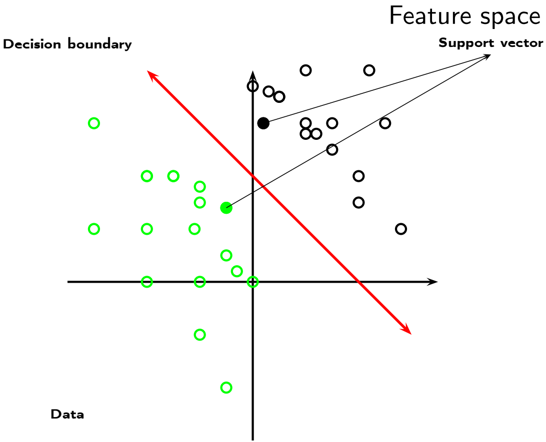
A diffusion kernel is a term coined by Laferty (2005) and it alludes to a Mercer kernel (or classifier in the context of Machine Learning), this results from solving the heat equation (diffusion equation) in the modeled manifold in the data set that have a known distribution (multinomial, gaussian, *q*-gaussian, etc.). In this short presentation the path that has been developed to obtain a diffusion kernel will be shown with the hypothesis that the data have a *q*-gaussian distribution with parameters (μ, σ).

Example Support Vector Machine (SVM)

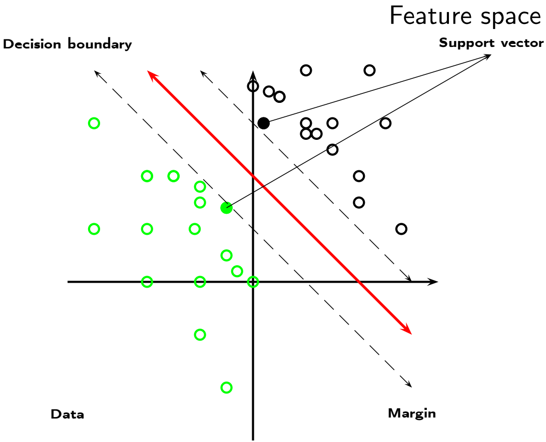
Feature space



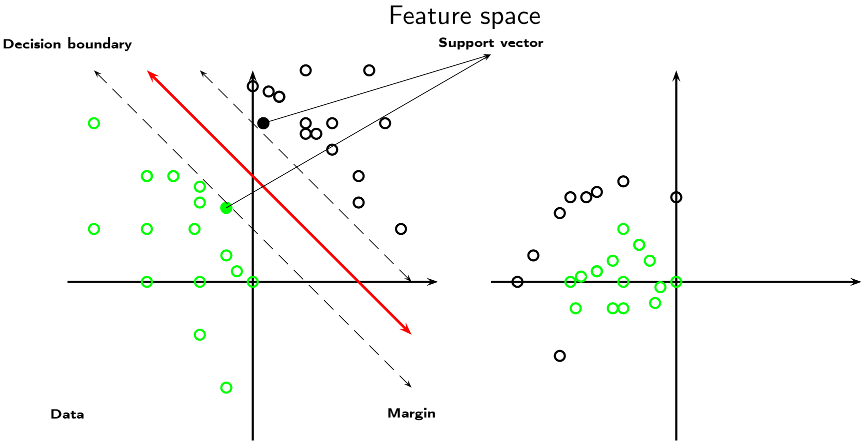
Example Support Vector Machine (SVM)



Example Support Vector Machine (SVM)



Example Support Vector Machine (SVM)



Example Support Vector Machine (SVM)

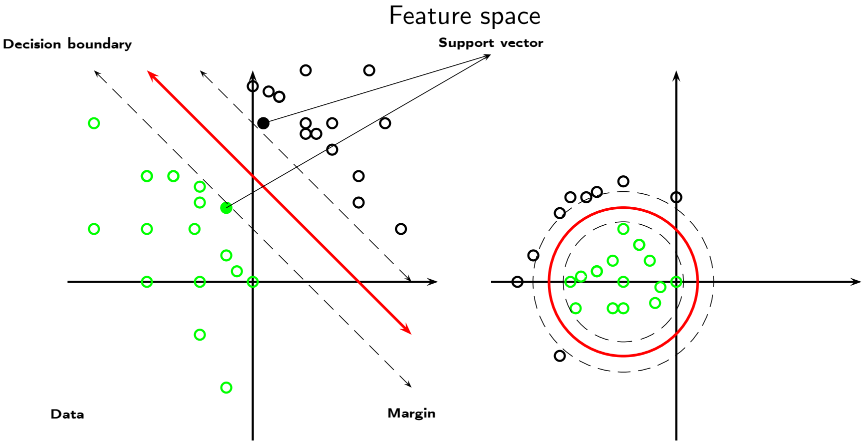


Figure 1: Ideas about the operation of SVM

Concepts, notations and equations of interest

- $p(x, \theta)$: Probability distribution for x a random variable in Ω y θ a vector of parameters in $\mathbb{R}^n$.

Concepts, notations and equations of interest

- $p(x, \theta)$: Probability distribution for x a random variable in Ω y θ a vector of parameters in $\mathbb{R}^n$.
- $\psi(\theta)$: Potential function, it results from writing the distribution $p(x, \theta)$ as $p(x, \boldsymbol{\theta}) = \exp(F(x) \cdot \boldsymbol{\theta} - \psi(\boldsymbol{\theta}))$ called **exponential family**, where $F(x) = (F_1(x), F_2(x) \ldots, F_n(x))$ and $\boldsymbol{\theta} = (\theta_1, \theta_2, \ldots, \theta_n)$.

Concepts, notations and equations of interest

- $\star$ $p(x, \theta)$: Probability distribution for x a random variable in Ω y θ a vector of parameters in $\mathbb{R}^n$.
- $\star$ $\psi(\theta)$: Potential function, it results from writing the distribution $p(x, \theta)$ as $p(x, \boldsymbol{\theta}) = \exp(F(x) \cdot \boldsymbol{\theta} - \psi(\boldsymbol{\theta}))$ called **exponential family**, where $F(x) = (F_1(x), F_2(x) \ldots, F_n(x))$ and $\boldsymbol{\theta} = (\theta_1, \theta_2, \ldots, \theta_n)$.
- $\star$ $E[f(x)]$: Expected value with respect to the distribution p for a function $f(x)$, is written

$$E[f(x)] = \int_\Omega f(x) p d\mu$$

Concepts, notations and equations of interest

- $p(x,\theta)$: Probability distribution for x a random variable in Ω y θ a vector of parameters in $\mathbb{R}^n$.
- $\psi(\theta)$: Potential function, it results from writing the distribution $p(x,\theta)$ as $p(x,\boldsymbol{\theta}) = \exp(F(x)\cdot\boldsymbol{\theta} - \psi(\boldsymbol{\theta}))$ called **exponential family**, where $F(x) = (F_1(x), F_2(x)\ldots, F_n(x))$ and $\boldsymbol{\theta} = (\theta_1, \theta_2, \ldots, \theta_n)$.
- $E[f(x)]$: Expected value with respect to the distribution p for a function $f(x)$, is written

$$E[f(x)] = \int_\Omega f(x)p\,d\mu$$

- $\partial_i f(x)$: Partial derivative of $f(x)$ with respect to the i-th component of the vector θ, is written

$$\partial_i f(x) = \frac{\partial f(x)}{\partial\theta_i}\ .$$

Concepts, notations and equations of interest

- $p(x,\theta)$: Probability distribution for x a random variable in Ω y θ a vector of parameters in $\mathbb{R}^n$.
- $\psi(\theta)$: Potential function, it results from writing the distribution $p(x,\theta)$ as $p(x,\boldsymbol{\theta}) = \exp(F(x)\cdot\boldsymbol{\theta} - \psi(\boldsymbol{\theta}))$ called **exponential family**, where $F(x) = (F_1(x), F_2(x)\ldots, F_n(x))$ and $\boldsymbol{\theta} = (\theta_1, \theta_2, \ldots, \theta_n)$.
- $E[f(x)]$: Expected value with respect to the distribution p for a function $f(x)$, is written

$$E[f(x)] = \int_\Omega f(x) p \, d\mu$$

- $\partial_i f(x)$: Partial derivative of $f(x)$ with respect to the i-th component of the vector θ, is written

$$\partial_i f(x) = \frac{\partial f(x)}{\partial \theta_i} \, .$$

- $\ell = \log p$: Score function, logarithm of the probability distribution.

Concepts, notations and equations of interest

- ⋆ g_{ij}^{F}: Components of Fisher's metric, defined as

$$g_{ij}^{F} = \int_{\Omega} (\partial_i \ell)(\partial_j \ell)\, p d\mu \ .$$

Concepts, notations and equations of interest

⋆ g_{ij}^F: Components of Fisher's metric, defined as

$$g_{ij}^F = \int_\Omega (\partial_i \ell)(\partial_j \ell)\, p d\mu \; .$$

⋆ $\frac{\partial f}{\partial t} = \Delta_g f$: Heat equation or diffusion equation where Δ is the operator Laplace - Beltrami defined in terms of the metric as

$$\Delta_g f = \frac{1}{\sqrt{\det g}} \sum_j \frac{\partial}{\partial x_j} \left(\sum_i g^{ij} \sqrt{\det g} \frac{\partial f}{\partial x_i} \right) \; ,$$

g^{ij} are the components of the inverse of the metric $g = [g_{ij}]$.

Concepts, notations and equations of interest

⋆ g^F_{ij}: Components of Fisher's metric, defined as

$$g^F_{ij} = \int_\Omega (\partial_i \ell)(\partial_j \ell)\, p d\mu \; .$$

⋆ $\frac{\partial f}{\partial t} = \Delta_g f$: Heat equation or diffusion equation where Δ is the operator Laplace - Beltrami defined in terms of the metric as

$$\Delta_g f = \frac{1}{\sqrt{\det g}} \sum_j \frac{\partial}{\partial x_j} \left(\sum_i g^{ij} \sqrt{\det g} \frac{\partial f}{\partial x_i} \right) \; ,$$

g^{ij} are the components of the inverse of the metric $g = [g_{ij}]$.

⋆ $\Gamma_{ij,k}$: Christofell symbols, defined as

$$\Gamma_{ij,k} = \sum_{h=1}^{n} \frac{1}{2} \left[\partial_i g_{jh} + \partial_j g_{ih} - \partial_h g_{ij} \right] g^{hk} \; .$$

- $\star$ R^l_{ijk}: Components of the metric tensor, are calculated by means of the expression

$$R^l_{ijk} = \sum_{h=1}^{n} \left[\Gamma^h_{ik} \Gamma^l_{jh} - \Gamma^h_{jk} \Gamma^l_{ih} \right] + \partial_j \Gamma^l_{ik} - \partial_i \Gamma^l_{jk} \ .$$

- ⋆ R^l_{ijk}: Components of the metric tensor, are calculated by means of the expression

$$R^l_{ijk} = \sum_{h=1}^{n} \left[\Gamma^h_{ik}\Gamma^l_{jh} - \Gamma^h_{jk}\Gamma^l_{ih}\right] + \partial_j\Gamma^l_{ik} - \partial_i\Gamma^l_{jk} \; .$$

- ⋆ Geodesic curve: It is obtained by solving the system of homogeneous second order differential equations

$$\frac{d\theta_k}{dt} + \sum_{i,j=1}^{n} \Gamma_{ij,k}\frac{d\theta_i}{dt}\frac{d\theta_j}{dt}$$

where each θ_i are the components of the parameter θ.

- ⋆ R_{ijk}^{l}: Components of the metric tensor, are calculated by means of the expression

$$R_{ijk}^{l} = \sum_{h=1}^{n} \left[\Gamma_{ik}^{h} \Gamma_{jh}^{l} - \Gamma_{jk}^{h} \Gamma_{ih}^{l} \right] + \partial_j \Gamma_{ik}^{l} - \partial_i \Gamma_{jk}^{l} \; .$$

- ⋆ Geodesic curve: It is obtained by solving the system of homogeneous second order differential equations

$$\frac{d\theta_k}{dt} + \sum_{i,j=1}^{n} \Gamma_{ij,k} \frac{d\theta_i}{dt} \frac{d\theta_j}{dt}$$

where each θ_i are the components of the parameter θ.

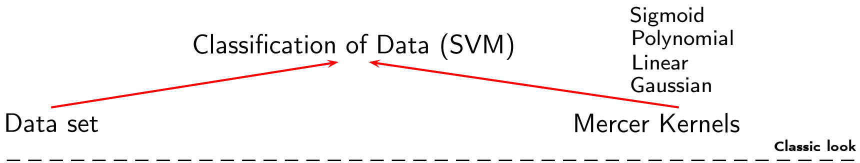
- ⋆ ρ: Geodesic distance, parametrizing the geodetic curve as $\gamma(t)$, this distance is

$$\rho = \int_a^b \sqrt{g_\gamma \left(\dot{\gamma}, \dot{\gamma} \right)} dt$$

where $\dot{\gamma}$ is the derivate of γ with respect to t.

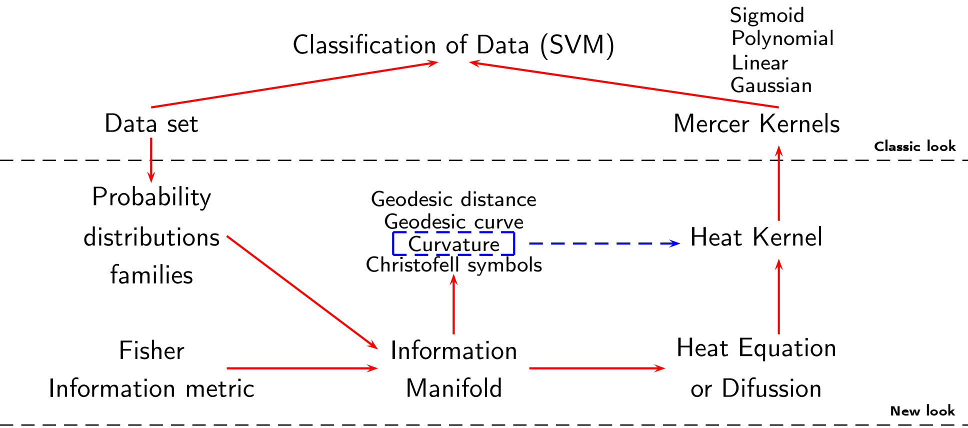
Way to obtain a diffusion kernel

Lafferty and Lebanon propose in their article *Diffusion Kernels on Statistical Manifold* (Laferty and Lebanon, 2005) the following way of work



Way to obtain a diffusion kernel

Lafferty and Lebanon propose in their article *Diffusion Kernels on Statistical Manifold* (Laferty and Lebanon, 2005) the following way of work



Heat kernel by Grigor'yan and Noguchi

1. In the case of the Euclidean space $\mathbb{R}^n$, the Heat Kernel is given by

$$K_t(x,y) = \frac{1}{(4\pi t)^{n/2}} \exp\left(-\frac{\|x-y\|^2}{4t}\right) = \frac{1}{(4\pi t)^{n/2}} \exp\left(-\frac{d^2(x,y)}{4t}\right)$$

where $\|x-y\|^2$ is the square of the Euclidean distance (geodesic distance) between points x and y.

Heat kernel by Grigor'yan and Noguchi

1. In the case of the Euclidean space $\mathbb{R}^n$, the Heat Kernel is given by

$$K_t(x,y) = \frac{1}{(4\pi t)^{n/2}} \exp\left(-\frac{\|x-y\|^2}{4t}\right) = \frac{1}{(4\pi t)^{n/2}} \exp\left(-\frac{d^2(x,y)}{4t}\right)$$

where $\|x-y\|^2$ is the square of the Euclidean distance (geodesic distance) between points x and y.

2. On the hyperbolic space $\mathbb{H}^n$, the heat kernel is given by

$$K_t(x,x') = \begin{cases} \frac{(-1)^m}{(2\pi)^m} \frac{1}{\sqrt{4\pi t}} \left(\frac{1}{\sinh\rho}\frac{\partial}{\partial\rho}\right)^m \exp\left(-m^2 t - \frac{\rho^2}{4t}\right) & \text{If } n = 2m+1 \\ \frac{(-1)^m}{(2\pi)^m} \frac{\sqrt{2}}{\sqrt{(4\pi t)^3}} \left(\frac{1}{\sinh\rho}\frac{\partial}{\partial\rho}\right)^m \int_\rho^\infty \frac{s\exp\left(-\frac{(2m+1)^2 t}{4} - \frac{s^2}{4t}\right)}{\sqrt{\cosh s - \cosh\rho}} ds & \text{If } n = 2m+2 \end{cases}$$

where $\rho = d(x,x')$ is the geodesic distance between the two points in the plane $\mathbb{H}^n$. If $n=2$ ($m=0$ in the second case) then

$$K_t(x,x') = \frac{\sqrt{2}}{(4\pi t)^{\frac{3}{2}}} \exp\left(-\frac{t}{4}\right) \int_\rho^\infty \frac{s\exp\left(-\frac{s^2}{4t}\right)}{\sqrt{\cosh s - \cosh\rho}} ds \, .$$

Tsallis entropy

In the context of non-extensive statistical mechanics, Constantino Tsallis (in 1988) defines entropy relative to q as

$$S_q = \frac{1}{1-q}\left(\sum_i p_i^q - 1\right) = \frac{1}{1-q}\left(h_q - 1\right)$$

where $\sum_i p(x_i) = \sum_i p_i = 1$, q is a fixed value less than 3 called **entropy index** and h_q is the functional $h_q = E[p^q]$ ($E[\cdot]$ it can be summation or integral) that allows defining an expected value relative to q. So, if $q \to 1$ the Shannon entropy

$$S = -\sum_i p\left(x_i\right) \log p\left(x_i\right)$$

used in classical statistical mechanics is obtained.

The description makes sense when defining a pair of inverse functions one of the other, called q-exponential and q-logarithm that generalize the exponential and the logarithm, recovering these when q tends to 1.

The q-exponential function

The q-exponential function is defined as

$$\exp_q(x) = [1 + (1-q)x]_+^{\frac{1}{1-q}}$$

for $-\infty < q < 3$. The derivative for a fixed q value is

$$\frac{d}{dx}\exp_q(x) = \left[\exp_q(x)\right]^q .$$

The q-exponential function

The q-exponential function is defined as

$$\exp_q(x) = [1 + (1 - q)x]_+^{\frac{1}{1-q}}$$

for $-\infty < q < 3$. The derivative for a fixed q value is

$$\frac{d}{dx}\exp_q(x) = \left[\exp_q(x)\right]^q \ .$$

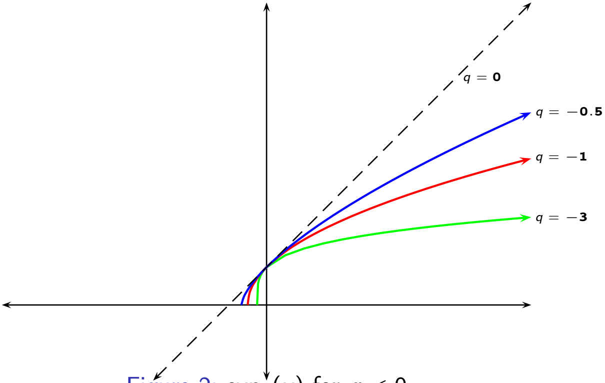


Figure 2: $\exp_q(x)$ for $q < 0$

Graphs for the function q-exponential

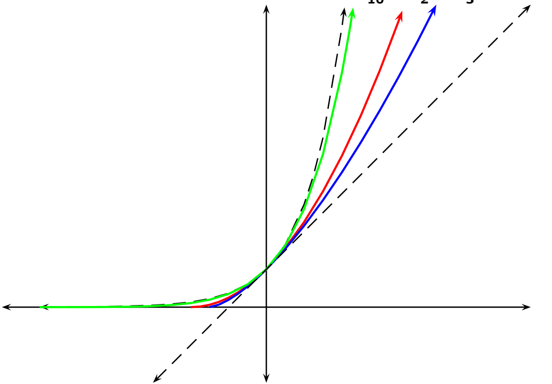


Figure 3: $\exp_q(x)$ for $0 < q < 1$

Graphs for the function q-exponential

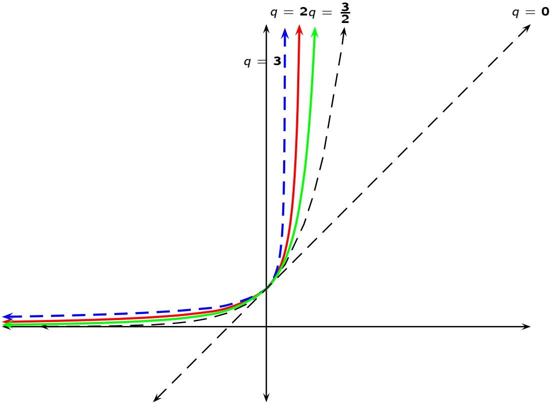


Figure 4: $\exp_q(x)$ for $1 < q < 3$

The q-logarithm function

The inverse of the q-exponential function, the q-logarithm, is given by

$$\ln_q x = \frac{x^{1-q} - 1}{1 - q}$$

provided that $x > 0$. The graph for some values of q is presented below, as well as its derivative for q fixed.

The q-logarithm function

The inverse of the q-exponential function, the q-logarithm, is given by

$$\ln_q x = \frac{x^{1-q} - 1}{1 - q}$$

provided that $x > 0$. The graph for some values of q is presented below, as well as its derivative for q fixed.

$\frac{d}{dx}[\log_q(x)] = \frac{1}{x^q}$

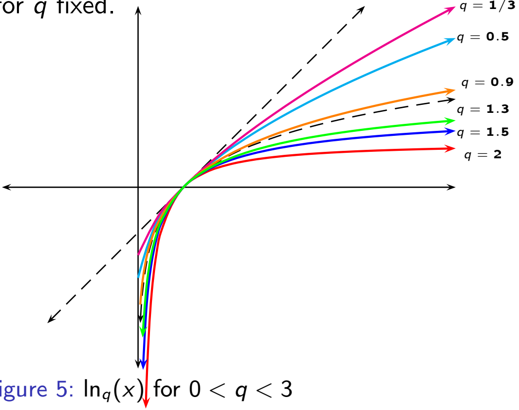


Figure 5: $\ln_q(x)$ for $0 < q < 3$

q-Gaussian Distribution

The q-gaussian distribution has density function

$$p_q(x,\theta) = \frac{1}{Z_{q,\sigma}} \exp_q\left(-\frac{(x-\mu)^2}{(3-q)\sigma^2}\right) = \frac{1}{Z_{q,\sigma}}\left[1 - \frac{(1-q)}{(3-q)}\frac{(x-\mu)^2}{\sigma^2}\right]^{\frac{1}{1-q}}$$

where $\theta = (\mu, \sigma)$ are the parameters on which the manifold of information is defined, $Z_{q,\sigma}$ is the normalization constant that depends on q, is written as $Z_{q,\sigma} = A_q\sigma$.

q-Gaussian Distribution

The q-gaussian distribution has density function

$$p_q(x,\theta) = \frac{1}{Z_{q,\sigma}} \exp_q\left(-\frac{(x-\mu)^2}{(3-q)\sigma^2}\right) = \frac{1}{Z_{q,\sigma}}\left[1 - \frac{(1-q)}{(3-q)}\frac{(x-\mu)^2}{\sigma^2}\right]^{\frac{1}{1-q}}$$

where $\theta = (\mu, \sigma)$ are the parameters on which the manifold of information is defined, $Z_{q,\sigma}$ is the normalization constant that depends on q, is written as $Z_{q,\sigma} = A_q \sigma$.

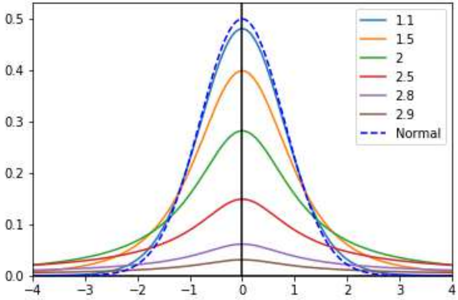


Figure 6: q-gaussian distribution

Normalization constant

The normalization constant is obtained by satisfying the expression

$$\int_{-\infty}^{\infty} f(x)dx = 1 \text{ or } Z_{q,\sigma} = \int_{-\infty}^{\infty} \left[1 - \frac{(1-q)}{(3-q)}\frac{(x-\mu)^2}{\sigma^2}\right]^{\frac{1}{1-q}} dx \ .$$

Normalization constant

The normalization constant is obtained by satisfying the expression

$$\int_{-\infty}^{\infty} f(x)dx = 1 \text{ or } Z_{q,\sigma} = \int_{-\infty}^{\infty} \left[1 - \frac{(1-q)}{(3-q)}\frac{(x-\mu)^2}{\sigma^2}\right]^{\frac{1}{1-q}} dx \ .$$

By means of a variable change, the constant A_q is defined in terms of the Beta function (or the Gamma function) for some values of q

1. $A_q = \sqrt{\frac{3-q}{1-q}} B\left(\frac{2-q}{1-q}, \frac{1}{2}\right)$ if $-\infty < q < 1$. In this situation the admissible domain for x is $\left[-\frac{\sigma}{\sqrt{1-q}}, \frac{\sigma}{\sqrt{1-q}}\right]$.
2. $A_q = \sqrt{2\pi}$ if $q = 1$.
3. $A_q = \sqrt{\frac{3-q}{q-1}} B\left(\frac{3-q}{2(q-1)}, \frac{1}{2}\right)$ if $1 < q < 3$. The domain for x are all real numbers.

Particular cases

1. Gaussian distribution ($q = 1$).
2. Cauchy distribution ($q = 2$).
3. t-Students distribution ($q = 1 + \frac{2}{n+1}$ with $n \in \mathbb{N}$).
4. Uniform distribution ($q \to -\infty$).
5. Wigner semicircle distribution ($q = -1$).

The q-gaussian distribution belong an exponential family

According to the definition of the function q-logarithm applied to the q-gaussian distribution it is possible to write

$$\log_q p_q = \frac{1}{1-q}\left(\left(\frac{1}{Z_{q,\sigma}}\exp_q\left(-\frac{(x-\mu)^2}{(3-q)\sigma^2}\right)\right)^{1-q} - 1\right) ,$$

The q-gaussian distribution belong an exponential family

According to the definition of the function q-logarithm applied to the q-gaussian distribution it is possible to write

$$\begin{aligned}
\log_q p_q &= \frac{1}{1-q}\left(\left(\frac{1}{Z_{q,\sigma}}\exp_q\left(-\frac{(x-\mu)^2}{(3-q)\sigma^2}\right)\right)^{1-q} - 1\right) , \\
&= \underbrace{\frac{Z_{q,\sigma}^{q-1}}{3-q}\frac{2\mu}{\sigma^2}}_{\theta_1}\underbrace{x}_{F_1(x)} + \underbrace{\frac{Z_{q,\sigma}^{q-1}}{3-q}\frac{1}{\sigma^2}}_{\theta_2}\underbrace{(-x^2)}_{F_2(x)} - \underbrace{\left[\frac{Z_{q,\sigma}^{q-1}}{3-q}\frac{\mu^2}{\sigma^2} - \log_q\left(\frac{1}{Z_{q,\sigma}}\right)\right]}_{\psi_q(\mu,\sigma)} , \\
&= \theta_1 F_1(x) + \theta_2 F_2(x) - \psi_q(\mu,\sigma) .
\end{aligned}$$

The q-gaussian distribution belong an exponential family

According to the definition of the function q-logarithm applied to the q-gaussian distribution it is possible to write

$$\begin{aligned}\log_q p_q &= \frac{1}{1-q}\left(\left(\frac{1}{Z_{q,\sigma}}\exp_q\left(-\frac{(x-\mu)^2}{(3-q)\sigma^2}\right)\right)^{1-q}-1\right) ,\\ &= \underbrace{\frac{Z_{q,\sigma}^{q-1}}{3-q}\frac{2\mu}{\sigma^2}}_{\theta_1}\underbrace{x}_{F_1(x)}+\underbrace{\frac{Z_{q,\sigma}^{q-1}}{3-q}\frac{1}{\sigma^2}}_{\theta_2}\underbrace{(-x^2)}_{F_2(x)}-\underbrace{\left[\frac{Z_{q,\sigma}^{q-1}}{3-q}\frac{\mu^2}{\sigma^2}-\log_q\left(\frac{1}{Z_{q,\sigma}}\right)\right]}_{\psi_q(\mu,\sigma)} ,\\ &= \theta_1 F_1(x)+\theta_2 F_2(x)-\psi_q(\mu,\sigma) .\end{aligned}$$

Then the q-gaussian distribution is an element in the family q-exponential with parameters and function q-potential

$$\theta_1 = \frac{Z_{q,\sigma}^{q-1}}{3-q}\frac{2\mu}{\sigma^2} , \qquad \theta_2 = -\frac{Z_{q,\sigma}^{q-1}}{3-q}\frac{1}{\sigma^2} ,$$

$$\psi_q(\theta_1,\theta_2) = -\frac{\theta_1^2}{4\theta_2}-\log_q\left[(-d_q\,\theta_2)^{\frac{1}{3-q}}\right] , \qquad \text{with } d_q = \frac{3-q}{A_q^2} .$$

Fisher's metrics and its representations

On the manifold defined by the q-gaussian distribution, it is possible to define two metrics. To do this, Amari (2009) defines the functional Ω_q for a probability distribution p as

$$\Omega_{q,p} = \int_{\Omega} p^q d\mu \ ,$$

Fisher's metrics and its representations

On the manifold defined by the q-gaussian distribution, it is possible to define two metrics. To do this, Amari (2009) defines the functional Ω_q for a probability distribution p as

$$\Omega_{q,p} = \int_\Omega p^q d\mu \ ,$$

it allows to define a probability distribution q-relative

$$\hat{p}_q = \frac{1}{\Omega_{p,q}} p^q \qquad \text{where} \qquad \int_\Omega \hat{p} d\mu = 1 \ .$$

Fisher's metrics and its representations

On the manifold defined by the q-gaussian distribution, it is possible to define two metrics. To do this, Amari (2009) defines the functional Ω_q for a probability distribution p as

$$\Omega_{q,p} = \int_{\Omega} p^q d\mu \ ,$$

it allows to define a probability distribution q-relative

$$\hat{p}_q = \frac{1}{\Omega_{p,q}} p^q \qquad \text{where} \qquad \int_{\Omega} \hat{p} d\mu = 1 \ .$$

along with the q-expectation

$$E[f(x)] = \int_{\Omega} f(x) \hat{p}_q d\mu = \frac{1}{\Omega_{q,p}} \int_{\Omega} f(x) p^q d\mu \ .$$

For the q-gaussian distribution the relation is fulfilled (Tanaya, 2011)

$$\Omega_{q,p} = \frac{3-q}{2} Z_{q,p}^{1-q} = \frac{3-q}{2} A_q^{1-q} \sigma^{1-q} \ .$$

Fisher's metrics and its representations

One of the metrics defined in the manifold is the usual Fisher g_{ij}^{F} induced by the Score function $\ell = \log p_q$ and the other is the q-Fisher's metric defined by Amari (2009) and what can be written about the distribution $\hat{p}$ as

$$g_{ij}^{(q)} = E_{\hat{p}}\left[(\partial_i \ell_q)(\partial_j \ell_q)\, q p^{q-1}\right] = \frac{q}{\Omega_{q,p}} \int_{\Omega} (\partial_i \ell_q)(\partial_j \ell_q)\, p^{2q-1} d\mu \,.$$

where $\ell_q = \log_q p_q$.

Fisher's metrics and its representations

One of the metrics defined in the manifold is the usual Fisher g_{ij}^{F} induced by the Score function $\ell = \log p_q$ and the other is the q-Fisher's metric defined by Amari (2009) and what can be written about the distribution $\hat{p}$ as

$$g_{ij}^{(q)} = E_{\hat{p}}\left[\left(\partial_i \ell_q\right)\left(\partial_j \ell_q\right) q p^{q-1}\right] = \frac{q}{\Omega_{q,p}} \int_{\Omega} \left(\partial_i \ell_q\right)\left(\partial_j \ell_q\right) p^{2q-1} d\mu \ .$$

where $\ell_q = \log_q p_q$. The two metrics are related by means of equality (Amari 2009)

$$g_{ij}^{(q)} = \frac{q}{\Omega_{q,p}} g_{ij}^{F} \ .$$

Fisher's metrics and its representations

One of the metrics defined in the manifold is the usual Fisher g_{ij}^F induced by the Score function $\ell = \log p_q$ and the other is the q-Fisher's metric defined by Amari (2009) and what can be written about the distribution $\hat{p}$ as

$$g_{ij}^{(q)} = E_{\hat{p}}\left[(\partial_i \ell_q)(\partial_j \ell_q)\, q p^{q-1}\right] = \frac{q}{\Omega_{q,p}} \int_\Omega (\partial_i \ell_q)(\partial_j \ell_q)\, p^{2q-1} d\mu \,.$$

where $\ell_q = \log_q p_q$. The two metrics are related by means of equality (Amari 2009)

$$g_{ij}^{(q)} = \frac{q}{\Omega_{q,p}} g_{ij}^F \,.$$

It is also shown that Fisher's q-metric is of the form

$$g_{ij}^{(q)} = \partial_i \partial_j \psi_q$$

which induces a Hessian manifold. Deriving the function q-potential in terms of parameters θ_1 and θ_2 we get the matrix $g^{(q)}$ and by a change of coordinates it is possible to obtain a matrix diagonal $g_*^{(q)}$.

q-Fisher's metrics and its matrix representations

$$\psi_q(\theta_1,\theta_2) = -\frac{\theta_1^2}{4\theta_2} - \log_q\left[(-d_q\ \theta_2)^{\frac{1}{3-q}}\right]$$

Coordinates (θ_1,θ_2)	Coordinates (μ,σ)
$g^{(q)} = \begin{bmatrix} \frac{-1}{2\theta_2} & \frac{\theta_1}{2\theta_2^2} \\ \frac{\theta_1}{2\theta_2^2} & -\frac{\theta_1^2}{2\theta_2^3} + \frac{1}{3-q}\frac{\Omega_{q,\theta_2}^{-1}}{\theta_2^2} \end{bmatrix}$	$g_*^{(q)} = \begin{bmatrix} \frac{\Omega_{q,\sigma}^{-1}}{\sigma^2} & 0 \\ 0 & \frac{(3-q)\Omega_{q,\sigma}^{-1}}{\sigma^2} \end{bmatrix}$
$\det\left(g^{(q)}\right) = \frac{1}{2(3-q)}\frac{\Omega_{q,\theta_2}^{-1}}{(-\theta_2)^3}$	$\det\left(g_*^{(q)}\right) = \frac{(3-q)\Omega_{q,\sigma}}{\sigma^4}$
$\left[g^{(q)}\right]^{-1} = \begin{bmatrix} (3-q)\Omega\theta_1^2 - 2\theta_2 & (3-q)\Omega\theta_1\theta_2 \\ (3-q)\Omega\theta_1\theta_2 & (3-q)\Omega\theta_2^2 \end{bmatrix}$	$\left[g_*^{(q)}\right]^{-1} = \begin{bmatrix} \Omega\sigma^2 & 0 \\ 0 & \frac{\Omega\sigma^2}{3-q} \end{bmatrix}$

Fisher's metrics and its matrix representations

$$\psi_q(\theta_1,\theta_2) = -\frac{\theta_1^2}{4\theta_2} - \log_q\left[(-d_q\,\theta_2)^{\frac{1}{3-q}}\right]$$

$$g_{ij}^F = \frac{\Omega_{q,p}}{q} g_{ij}^{(q)}$$

Coordinates (θ_1,θ_2)	Coordinates (μ,σ)
$g^F = \begin{bmatrix} \frac{-\Omega}{2q\theta_2} & \frac{\Omega\theta_1}{2q\theta_2^2} \\ \frac{\Omega\theta_1}{2q\theta_2^2} & -\frac{\Omega\theta_1^2}{2q\theta_2^3} + \frac{1}{q(3-q)}\frac{1}{\theta_2^2} \end{bmatrix}$	$g_*^F = \begin{bmatrix} \frac{1}{q\sigma^2} & 0 \\ 0 & \frac{3-q}{q\sigma^2} \end{bmatrix}$
$\det\left(g^{(q)}\right) = \frac{1}{2q^2(3-q)}\frac{\Omega}{(-\theta_2)^3}$	$\det\left(g_*^F\right) = \frac{3-q}{q^2\sigma^4}$
$\left[g^F\right]^{-1} = \begin{bmatrix} (3-q)q\theta_1^2 - \frac{2q\theta_2}{\Omega} & (3-q)q\theta_1\theta_2 \\ (3-q)q\theta_1\theta_2 & (3-q)q\theta_2^2 \end{bmatrix}$	$\left[g_*^F\right]^{-1} = \begin{bmatrix} q\sigma^2 & 0 \\ 0 & \frac{q}{3-q}\sigma^2 \end{bmatrix}$

Christoffel symbols and curvature

Deriving the components of the matrix g_*^F regarding the parameters (μ, σ), it is possible to obtain the Christoffel symbols as summarized in continuation

Derivadas de las componentes de la métrica			
$\partial_1 g_{11}^F = 0$ $\partial_2 g_{11}^F = -\frac{2}{q\sigma^3}$	$\partial_1 g_{12}^F = 0$ $\partial_2 g_{12}^F = 0$	$\partial_1 g_{21}^F = 0$ $\partial_2 g_{21}^F = 0$	$\partial_1 g_{22}^F = 0$ $\partial_2 g_{22}^F = -\frac{2(3-q)}{q\sigma^3}$

Christoffel symbols and curvature

Deriving the components of the matrix g_*^F regarding the parameters (μ, σ), it is possible to obtain the Christoffel symbols as summarized in continuation

Derivadas de las componentes de la métrica			
$\partial_1 g_{11}^F = 0$ $\partial_2 g_{11}^F = -\frac{2}{q\sigma^3}$	$\partial_1 g_{12}^F = 0$ $\partial_2 g_{12}^F = 0$	$\partial_1 g_{21}^F = 0$ $\partial_2 g_{21}^F = 0$	$\partial_1 g_{22}^F = 0$ $\partial_2 g_{22}^F = -\frac{2(3-q)}{q\sigma^3}$

Christoffel symbols			
$\Gamma_{11,1}^F = 0$ $\Gamma_{21,1}^F = -\frac{1}{\sigma}$	$\Gamma_{11,2}^F = \frac{1}{(3-q)\sigma}$ $\Gamma_{21,2}^F = 0$	$\Gamma_{12,1}^F = -\frac{1}{\sigma}$ $\Gamma_{22,1}^F = 0$	$\Gamma_{12,2}^F = 0$ $\Gamma_{22,2}^F = -\frac{1}{\sigma}$

Christoffel symbols and curvature

Deriving the components of the matrix g_*^F regarding the parameters (μ, σ), it is possible to obtain the Christoffel symbols as summarized in continuation

Derivadas de las componentes de la métrica			
$\partial_1 g_{11}^F = 0$ $\partial_2 g_{11}^F = -\frac{2}{q\sigma^3}$	$\partial_1 g_{12}^F = 0$ $\partial_2 g_{12}^F = 0$	$\partial_1 g_{21}^F = 0$ $\partial_2 g_{21}^F = 0$	$\partial_1 g_{22}^F = 0$ $\partial_2 g_{22}^F = -\frac{2(3-q)}{q\sigma^3}$

Christoffel symbols			
$\Gamma_{11,1}^F = 0$ $\Gamma_{21,1}^F = -\frac{1}{\sigma}$	$\Gamma_{11,2}^F = \frac{1}{(3-q)\sigma}$ $\Gamma_{21,2}^F = 0$	$\Gamma_{12,1}^F = -\frac{1}{\sigma}$ $\Gamma_{22,1}^F = 0$	$\Gamma_{12,2}^F = 0$ $\Gamma_{22,2}^F = -\frac{1}{\sigma}$

Metric Tensor		
$R_{212}^1 = \frac{1}{\sigma^2}$	$R_{212}^2 = 0$	$R_{1212} = g_{11} R_{212}^1 = -\frac{1}{q\sigma^4}$

Curvature and Geometry	
$k = \frac{R_{1212}}{\det(g_*^F)} = -\frac{q}{3-q} < 0$	Negative constant curvature (hyperbolic space)

Geodesic curves

Assuming that the coordinates (μ, σ) can be parametric depending on t and with the Christoffel symbols previously found, it is possible to define a system of homogeneous second order differential equations that describes the geodesic curves for the hyperbolic manifold generated by the q-gaussian distribution

$$\frac{d^2\mu}{dt^2} - \frac{2}{\sigma}\frac{d\mu}{dt}\frac{d\sigma}{dt} = 0$$
$$\frac{d^2\sigma}{dt^2} + \frac{1}{(3-q)\sigma}\left(\frac{d\mu}{dt}\right)^2 - \frac{1}{\sigma}\left(\frac{d\sigma}{dt}\right)^2 = 0 \, .$$

Geodesic curves

Assuming that the coordinates (μ, σ) can be parametric depending on t and with the Christoffel symbols previously found, it is possible to define a system of homogeneous second order differential equations that describes the geodesic curves for the hyperbolic manifold generated by the q-gaussian distribution

$$\frac{d^2\mu}{dt^2} - \frac{2}{\sigma}\frac{d\mu}{dt}\frac{d\sigma}{dt} = 0$$
$$\frac{d^2\sigma}{dt^2} + \frac{1}{(3-q)\sigma}\left(\frac{d\mu}{dt}\right)^2 - \frac{1}{\sigma}\left(\frac{d\sigma}{dt}\right)^2 = 0\,.$$

With the substitution $w = \frac{d\mu}{d\sigma}$ it is possible to show that the curve that solves this system is

Geodesic curves

Assuming that the coordinates (μ, σ) can be parametric depending on t and with the Christoffel symbols previously found, it is possible to define a system of homogeneous second order differential equations that describes the geodesic curves for the hyperbolic manifold generated by the q-gaussian distribution

$$\frac{d^2\mu}{dt^2} - \frac{2}{\sigma}\frac{d\mu}{dt}\frac{d\sigma}{dt} = 0$$

$$\frac{d^2\sigma}{dt^2} + \frac{1}{(3-q)\sigma}\left(\frac{d\mu}{dt}\right)^2 - \frac{1}{\sigma}\left(\frac{d\sigma}{dt}\right)^2 = 0 \, .$$

With the substitution $w = \frac{d\mu}{d\sigma}$ it is possible to show that the curve that solves this system is

$$(\mu - h)^2 + (3-q)\sigma^2 = \frac{3-q}{k^2}$$

where h and k are constants that possibly depend on q. This curve is an ellipse with center in $(h, 0)$, that is, on the axis μ. If $q = 2$ (Cauchy distribution) the curves are circumferences of radio $\frac{1}{k}$.

Further works

- ⋆ Find the geodesic distance for a q-gaussiana distribution for any $1 < q < 3$.
- ⋆ The Box-Muller method is applicable for q-gaussian distribution (Thistleton, 2007) generating random data with this distribution.
- ⋆ Program in Python these diffusion kernels for the manifold generated by the q-gaussian distribution.
- ⋆ Define appropriate Christoffel symbols for the q-metric that allow me to find the curvature for the system (μ, σ) and that is in accordance with the result $k = \frac{5-3q}{(q-3)^2(2q-3)}$ (Matsuzoe, 2014).
- ⋆ Study another way to find distances by means of the heat equation (Keenan, 2013).

Bibliographic references

[1] Amari, Shun-ichi and Ohara, Atsumi. *Geometry of q-Exponential Family of Probability Distributions*. Entropy 13, 2011, pp. 1170-1185. doi: 10.3390/e13061170.

[2] Burns; Keith and Gidea; Marian. *Differential Geometry and Topology. With a View to Dynamic System*. Studies in advance mathematics. Chapman & Hall / CRC, 2005.

[3] Cadavid; Carlos and Vélez; Juan Diego. *A Remark on the Heat Equation and Minimal Morse Functions on Tori and Spheres*. Ingeniería y Ciencia, Vol. 09, No. 17, Enero-Junio 2013. EAFIT.

[4] Cardenas Montes; Miguel. *Support Vector Machine. Graphs, Statistics and Data Mining with Python*. Presentation, November of 2015. http://wwwae.ciemat.es/~cardenas/docs/curso_MD/svm.pdf

[5] Grigor'yan, Alexander and Noguchi, Masakazu. *The heat kernel on hyperbolic space*. Bulletin of the London Mathematical Society, 30, pp. 643-650, 1998.

[6] Keenan, Crane. *A New Approach to Computing Distance based on Heat Flow*. NY, 2013.

[7] Laferty; John and Lebanon; Guy. *Difussion Kernels on Statistical Manifolds*. Journal of Machine Learning Research, 6 (2005), pp. 129-163.

[8] Matsuzoe, Hiroshi. *Hessian structures on deformed exponential families and their conformal structures*. Differential Geometry and its Applications, 35 (2014), pp. 323-333.

[9] Matsuzoe, Hiroshi and Ohara, Atsumi. *Geometry for a q-exponential families*. WSPC - Proceedings, April 16, 2011.

[10] *1.4. Support Vector Machines*. Scikit Learn. http://scikit-learn.org/stable/modules/svm.html. September 23/ 2017.

[11] Tanaya, Daiki, Tanaka, Masaru and Matsuzoe, Hiroshi. *Notes on geometry of q-normal distributions*. WSPC - Proceedings, May 07, 2011.

[12] Thistleton, William and others. *Generalized Box-Müller Method for Generating q-Gaussian Random Deviates*. Journal IEEE Transsactions of Information Theory, volume 53, Issue 12, December 2007, pages 4805 - 4810.

Thank you!!