Hamilton-Jacobi approach to contrast functions and geodetical motion in information geometry

Juan Manuel Pérez-Pardo

uc3m | Universidad Carlos III de Madrid
Departamento de Matemáticas

o	j	f	n	z	g	p	s	c	g	t	b	l	j	z	t	s	l
F	l	o	r	i	o	M	C	i	a	g	l	i	a	c	u	a	m
r	d	j	e	i	x	a	d	s	p	b	i	o	t	n	u	b	q
a	n	g	i	m	u	r	x	t	e	n	s	a	u	g	n	w	r
n	i	w	x	b	l	c	u	w	m	l	P	t	S	z	a	h	m
c	q	j	G	a	i	o	y	n	n	l	a	a	t	v	o	w	x
o	y	o	i	g	s	L	l	y	z	s	t	b	e	m	l	g	t
V	g	w	u	b	l	a	p	F	c	m	r	t	f	q	f	x	g
e	j	r	s	j	h	u	k	a	l	b	i	t	a	z	u	c	i
n	r	w	e	s	t	d	n	b	v	f	z	h	n	t	m	v	b
t	e	f	p	b	F	a	b	i	o	D	i	C	o	s	m	o	m
r	z	z	p	y	n	t	m	o	k	m	a	j	M	d	r	p	w
i	p	a	e	p	w	o	i	M	t	x	V	z	a	c	c	t	d
g	t	v	M	b	x	l	f	e	x	g	i	o	n	l	z	y	o
l	q	k	a	e	w	p	s	l	n	c	t	h	c	v	k	k	m
i	b	n	r	r	t	h	t	e	c	p	a	p	i	d	b	s	h
a	k	s	m	d	r	q	m	j	n	b	l	y	n	b	z	c	i
d	q	D	o	m	e	n	i	c	o	F	e	l	i	c	e	m	a

o	j	f	n	z	g	p	s	c	g	t	b	l	j	z	t	s	l
F	l	o	r	i	o	M	C	i	a	g	l	i	a	c	u	a	m
r	d	j	e	i	x	a	d	s	p	b	i	o	t	n	u	b	q
a	n	g	i	m	u	r	x	t	e	n	s	a	u	g	n	w	r
n	i	w	x	b	l	c	u	w	m	l	P	t	S	z	a	h	m
c	q	j	G	a	i	o	y	n	n	l	a	a	t	v	o	w	x
o	y	o	i	g	s	L	l	y	z	s	t	b	e	m	l	g	t
V	g	w	u	b	l	a	p	F	c	m	r	t	f	q	f	x	g
e	j	r	s	j	h	u	k	a	l	b	i	t	a	z	u	c	i
n	r	w	e	s	t	d	n	b	v	f	z	h	n	t	m	v	b
t	e	f	p	b	F	a	b	i	o	D	i	C	o	s	m	o	m
r	z	z	p	y	n	t	m	o	k	m	a	j	M	d	r	p	w
i	p	a	e	p	w	o	i	M	t	x	V	z	a	c	c	t	d
g	t	v	M	b	x	l	f	e	x	g	i	o	n	l	z	y	o
l	q	k	a	e	w	p	s	l	n	c	t	h	c	v	k	k	m
i	b	n	r	r	t	h	t	e	c	p	a	p	i	d	b	s	h
a	k	s	m	d	r	q	m	j	n	b	l	y	n	b	z	c	i
d	q	D	o	m	e	n	i	c	o	F	e	l	i	c	e	m	a

Joint work with

F	l	o	r	i	o	M	C	i	a	g	l	i	a			
r						a										
a						r										
n						c					P		S			
c			G			o					a		t			
o			i			L					t		e			
V			u			a		F			r		f			
e			s			u		a			i		a			
n			e			d		b			z		n			
t			p		F	a	b	i	o	D	i	C	o	s	m	o
r			p			t		o			a		M			
i			e			o		M			V		a			
g			M					e			i		n			
l			a					l			t		c			
i			r					e			a		i			
a			m								l		n			
		D	o	m	e	n	i	c	o	F	e	l	i	c	e	

- Introduction and motivation
- Generating Potentials for arbitrary tensors
- Hamilton-Jacobi approach to contrast functions

- Pure states are a complex projective space

$$\tilde{h} = \frac{\langle d\psi \otimes d\psi \rangle}{\langle \psi | \psi \rangle} - \frac{\langle \psi | d\psi \rangle}{\langle \psi | \psi \rangle} \otimes \frac{\langle d\psi | \psi \rangle}{\langle \psi | \psi \rangle}$$

$$\pi : \mathcal{H} \to \mathrm{P}\mathcal{H}$$

$$\pi(|\psi\rangle) = \frac{|\psi\rangle\langle\psi|}{\langle \psi | \psi \rangle}$$

- Orthonormal basis $\{|e_j\rangle\}$ $\quad \langle e^j | e_k \rangle = \delta^j_k$

$$z_j = \langle e_j | \psi \rangle \qquad \bar{z}^j = \langle \psi | e^j \rangle$$

$$\tilde{h} = \frac{d\bar{z}^j \otimes dz_j}{\|z\|^2} - \frac{(\bar{z}^j dz_j \otimes) \otimes (d\bar{z}_k z^k)}{\|z\|^4}$$

- non-linear coordinates

$$z_j = \sqrt{p_j} e^{i\varphi_j} \qquad \left(p_j \geq 0 \,, \quad \sum_j p_j = 1 \right)$$

$$\tilde{h} = \frac{\mathrm{d}\bar{z}^j \otimes \mathrm{d}z_j}{\|z\|^2} - \frac{(\bar{z}^j \mathrm{d}z_j \otimes) \otimes (\mathrm{d}\bar{z}_k z^k)}{\|z\|^4}$$

$$\tilde{h} = \frac{1}{4} \sum_j \frac{1}{p_j} \mathrm{d}p_j \otimes \mathrm{d}p_j + \left[\mathrm{d}p_j \otimes \mathrm{d}\varphi_k + \mathrm{d}\varphi_j \otimes \mathrm{d}\varphi_k \right]$$

$$\mathcal{H} \equiv \mathcal{L}^2(\chi) \qquad |\psi\rangle \to \langle x\,|\psi\rangle = \psi(x)$$
$$\langle \psi\,|x\rangle = \psi^*(x)$$

- Immersion of a measure space manifold Θ into $\mathrm{P}\mathcal{H}$

$$\psi(x|\boldsymbol{\theta}) = \sqrt{p(x|\boldsymbol{\theta})}e^{i\alpha(x|\boldsymbol{\theta})}$$

$$\mathrm{d}\ln\psi(x|\boldsymbol{\theta}) = \frac{1}{2}\mathrm{d}\ln p(x|\boldsymbol{\theta}) + i\mathrm{d}\alpha(x|\boldsymbol{\theta})$$

$$h = g - iw$$
$$g = \tfrac{1}{4}E_p[(\mathrm{d}\ln p)\otimes(\mathrm{d}\ln p)] + E_p[\mathrm{d}\alpha\otimes\mathrm{d}\alpha] - E_p(\mathrm{d}\alpha)\otimes E_p(\mathrm{d}\alpha)$$
$$w = E_p[\mathrm{d}\ln p\wedge\mathrm{d}\alpha]$$

$$E_p(\mathrm{d}f) := \int_\chi (\mathrm{d}f)p\mathrm{d}x$$

$$\mathcal{F} = \tfrac{1}{4}E_p(\mathrm{d}\ln p\otimes\mathrm{d}\ln p)$$ Fisher-Rao metric

uc3m | Universidad Carlos III de Madrid
Departamento de Matemáticas

- Statistical Manifold: $\mathcal{M}$
  - Each point $p \in \mathcal{M}$ represents a probability density
- Fisher-Rao metric, a Riemannian metric: g
  - Infinitesimal "distinguishability" of probability densities
- Dualistic Structure: (g, ∇, ∇^*)

$$X, Y, Z \in \mathfrak{X}(\mathcal{M})$$

$$X(g(Y, Z)) = g(\nabla_X Y, Z) + g(Y, \nabla^*_X Z)$$

  - Usually defined by perturbation of the hermitean connection

$$\Gamma_{ijk} = {}_g\Gamma_{ijk} \pm T_{ijk}$$

- Statistical manifold: $\mathcal{M}$
- Divergence function: $S : \mathcal{M} \times \mathcal{M} \to \mathbb{R}$
  - $S(\mathbf{x}, \mathbf{y}) \geq 0$
  - $S(\mathbf{x}, \mathbf{y}) = 0 \iff \mathbf{x} = \mathbf{y}$
  - Oriented distance-like function

    $S(x, y) \neq S(y, x)$
  - If S is differentiable

$$\left.\frac{\partial S}{\partial x^i}\right|_{\mathbf{x}=\mathbf{y}} = \left.\frac{\partial S}{\partial y^i}\right|_{\mathbf{x}=\mathbf{y}} = 0$$

- S is a generating function
- Generates tensors intrinsically $\dfrac{\partial^2 S}{\partial x^i \partial x^j} = g_{ij}$

uc3m | Universidad Carlos III de Madrid
Departamento de Matemáticas

Outline

- Introduction and motivation
- Generating Potentials for arbitrary tensors
- Hamilton-Jacobi approach to contrast functions

uc3m | Universidad Carlos III de Madrid
Departamento de Matemáticas

- Some examples
- Hessian Riemannian Structures[1]:
  - $\mathbf{x} = \{x^i\}$ Local coordinates on $U \subset \mathcal{M}$
  - Convex function: $f : U \to \mathbb{R}$

$$g_{ij}(\mathbf{x}) := \frac{\partial^2 f}{\partial x^i \partial x^j}$$

This is not a tensor!

- Kähler Potential

$$\omega = \frac{\partial^2 f}{\partial z \partial \bar{z}}$$

Quote from Wikipedia: "There is no comparable way of describing a general Riemannian metric in terms of a single function."

[1]: J.J. Duistermaat. On Hessian Riemannian Structures. Asian J. Math, 5 (1) (2001), pp. 79-92

- Two copies of the manifold to define a two-point function

$$S : \mathcal{M} \times \mathcal{M} \to \mathbb{R}$$

- Provides several canonical structures:

$$X \in \mathfrak{X}(\mathcal{M})$$

$\pi_L : \mathcal{M} \times \mathcal{M} \to \mathcal{M}$ Left-lift: $X_L \in \mathfrak{X}(\mathcal{M} \times \mathcal{M})$

$\pi_R : \mathcal{M} \times \mathcal{M} \to \mathcal{M}$ Right-lift: $X_R \in \mathfrak{X}(\mathcal{M} \times \mathcal{M})$

Diagonal embedding:

$$d : \mathcal{M} \hookrightarrow \mathcal{M} \times \mathcal{M}$$
$$p \mapsto (p, p)$$

Restriction to the diagonal:

$$S|_d := d^* S$$

uc3m Universidad Carlos III de Madrid
Departamento de Matemáticas

- Consider the contrast function $S : \mathcal{M} \times \mathcal{M} \to \mathbb{R}$

$$(\mathcal{L}_{X_L}\mathcal{L}_{Y_L}S)\big|_d := g(X,Y)$$

  - It is clearly f-linear in X
  - Let us show that it is also f-linear in Y

$$\begin{aligned}\mathcal{L}_{X_L}\mathcal{L}_{(fY)_L}S &= \mathcal{L}_{X_L}(f\mathcal{L}_{Y_L}S) \\ &= f\mathcal{L}_{X_L}\mathcal{L}_{Y_L}S \quad + \; \mathcal{L}_{X_L}f \cdot \mathcal{L}_{Y_L}S\end{aligned}$$

- Consider the contrast function $S : \mathcal{M} \times \mathcal{M} \to \mathbb{R}$

$$(\mathcal{L}_{X_L}\mathcal{L}_{Y_L}S)\big|_d := g(X,Y)$$

  - It is clearly f-linear in X
  - Let us show that it is also f-linear in Y

$$\begin{aligned}(\mathcal{L}_{X_L}\mathcal{L}_{(fY)_L}S)\big|_d &= (\mathcal{L}_{X_L}(f\mathcal{L}_{Y_L}S))\big|_d \\ &= (f\mathcal{L}_{X_L}\mathcal{L}_{Y_L}S)\big|_d + (\mathcal{L}_{X_L}f\cdot\cancelto{0}{\mathcal{L}_{Y_L}S})\big|_d \\ &= fg(X,Y)\end{aligned}$$

Contrast Functions as Generating Potentials II

- Consider the contrast function $S : \mathcal{M} \times \mathcal{M} \to \mathbb{R}$

$$(\mathcal{L}_{X_L} \mathcal{L}_{Y_L} S)\big|_d := g(X, Y)$$

- It is clearly f-linear in X
- Let us show that it is also f-linear in Y

$$\begin{aligned}
(\mathcal{L}_{X_L} \mathcal{L}_{(fY)_L} S)\big|_d &= (\mathcal{L}_{X_L} (f \mathcal{L}_{Y_L} S))\big|_d \\
&= (f \mathcal{L}_{X_L} \mathcal{L}_{Y_L} S)\big|_d + (\mathcal{L}_{X_L} f \cdot \cancelto{0}{\mathcal{L}_{Y_L} S})\big|_d \\
&= f g(X, Y)
\end{aligned}$$

- One can define more rank 2 tensors:

$$(\mathcal{L}_{X_R} \mathcal{L}_{Y_R} S)\big|_d = (\mathcal{L}_{X_L} \mathcal{L}_{Y_L} S)\big|_d = -(\mathcal{L}_{X_R} \mathcal{L}_{Y_L} S)\big|_d = g(X, Y)$$

Contrast Functions as Generating Potentials III

- This procedure can be used for tensors of order 3
- One needs to find suitable linear combinations to cancel the non-tensorial terms

$$(\mathcal{L}_{X_L}\mathcal{L}_{Y_R}\mathcal{L}_{Z_R}S - \mathcal{L}_{X_R}\mathcal{L}_{Y_L}\mathcal{L}_{Z_L}S)\big|_d = T(X,Y,Z)$$

or

$$(\mathcal{L}_{X_L}\mathcal{L}_{Y_L}\mathcal{L}_{Z_R}S - \mathcal{L}_{X_R}\mathcal{L}_{Y_R}\mathcal{L}_{Z_L}S)\big|_d = T(X,Y,Z)$$

- There are 8 possible combinations that lead to tensorial quantities.
  - They define the same symmetric tensor up to a sign
  - In information geometry this is used to define pairs of dual connections

- Can one go to higher order tensors?
    - For instance the Riemann tensor

- Can one generate non-symmetric tensors?

uc3m | Universidad Carlos III de Madrid
Departamento de Matemáticas

- Introduction and motivation
- Generating Potentials for arbitrary tensors
- Hamilton-Jacobi approach to contrast functions

- Find a canonical contrast function
  - Given g T
  - Construct a two point function with vanishing derivatives
- Hamilton's Principal function

$$S[\gamma] = \int_{t_{\text{in}}}^{t_{\text{fin}}} \mathfrak{L}(\gamma(t), \dot{\gamma}(t)) \mathrm{d}t$$

  - Solution of the equations of motion γ
  - $\gamma(0) = \mathbf{x} \qquad \gamma(1) = \mathbf{y}$

$$S[\mathbf{x}, \mathbf{y}] = \int_0^1 \mathfrak{L}(\gamma(t), \dot{\gamma}(t)) \mathrm{d}t$$

  - $\mathrm{d}S = p_i(\mathbf{x}, \mathbf{y}) \mathrm{d}x^i - P_i(\mathbf{x}, \mathbf{y}) \mathrm{d}y^i$

$$\mathbf{x} = \mathbf{y} \qquad \gamma(t) = \mathbf{x} \qquad \Longrightarrow \qquad \begin{aligned} p_i(\mathbf{x}, \mathbf{y})|_{\mathbf{x}=\mathbf{y}} &= 0 \\ P_i(\mathbf{x}, \mathbf{y})|_{\mathbf{x}=\mathbf{y}} &= 0 \end{aligned}$$

uc3m | Universidad Carlos III de Madrid
Departamento de Matemáticas

Hamilton-Jacobi Approach to Contrast Functions II

- Choose the Lagrangian

$$\mathfrak{L}(x,v) = \frac{1}{2} g_{ij} v^i v^j + \frac{1}{6} T_{ijk} v^i v^j v^j$$

- Compute derivatives

$$\frac{\partial S}{\partial y^j} = g_{jk}(\mathbf{y}) v^k_{\text{fin}} + \frac{1}{2} T_{jkl}(\mathbf{y}) v^k_{\text{fin}} v^l_{\text{fin}}$$

  - We need the relation between $\mathbf{v}_{\text{fin}}$ and $\mathbf{x}, \mathbf{y}$
  - Taylor expansion

$$\gamma(t) = \gamma(1) + \mathbf{v}_{\text{fin}}(1-t) + \frac{1}{2}\dot{\mathbf{v}}_{\text{fin}}(1-t)^2 + \frac{1}{6}\ddot{\mathbf{v}}_{\text{fin}}(1-t)^3$$

$$\Longrightarrow \qquad \mathbf{v}_{\text{fin}} = \mathbf{y} - \mathbf{x} + \frac{1}{2}\dot{\mathbf{v}}_{\text{fin}} - \frac{1}{6}\ddot{\mathbf{v}}_{\text{fin}}$$

  - Eq. of motion

$$\dot{v}^l = -T^l_{jk} v^k \dot{v}^j - \Gamma^l_{jk} v^j v^k - \frac{1}{6} g^{lr} A_{rjks} v^j v^k v^s$$

- Assuming that $\dot{\mathbf{v}}$ is analytic in $\mathbf{v}$

$$\Delta^k := y^k - x^k$$

$$\frac{\partial S}{\partial y^j} = g(\mathbf{y})_{jk}\Delta^k - \frac{1}{2}\Gamma_{jkl}(\mathbf{y})\Delta^k\Delta^l + \frac{1}{2}T_{jkl}(\mathbf{y})\Delta^k\Delta^l + \frac{1}{6}\Gamma_{jkr}(\mathbf{y})\Gamma^r_{ls}(\mathbf{y})\Delta^k\Delta^l\Delta^s - \frac{1}{12}A_{jkls}(\mathbf{y})\Delta^k\Delta^l\Delta^s$$

$$\left.\frac{\partial^2 S}{\partial x^i \partial y^j}\right|_d = -g_{ij}(\mathbf{x}) \qquad \left.\frac{\partial^3 S}{\partial x^l \partial x^k \partial y^j}\right|_d = -\Gamma_{jkl} + T_{jkl}$$

By similar calculations: $$\left.\frac{\partial^3 S}{\partial y^l \partial y^k \partial x^j}\right|_d = -\Gamma_{jkl} - T_{jkl}$$

uc3m | Universidad Carlos III de Madrid
Departamento de Matemáticas

- Find a combination of 4th order derivatives that is f-linear

$$\mathcal{L}_{X_L}\mathcal{L}_{Y_R}\mathcal{L}_{Z_L}\mathcal{L}_{W_R}S\Big|_d =: LRLR$$

- Two possible linear combinations:

$$\begin{aligned} Q_1(X,Y,Z,W) =& (LRRL - RLLR) + (LRRR - RLLL) \\ &+ (LRLL - RLRR) + (LRLR - RLRL) \end{aligned}$$

$$\begin{aligned} Q_2(X,Y,Z,W) =& (LLLL - RRRR) + (LLLR - RRRL) \\ &+ (LLRL - RRLR) + (LLRR - RRLL) \end{aligned}$$

uc3m | Universidad **Carlos III** de Madrid
Departamento de Matemáticas

- Find a combination of 4th order derivatives that is f-linear

$$\mathcal{L}_{X_L}\mathcal{L}_{Y_R}\mathcal{L}_{Z_L}\mathcal{L}_{W_R}S\Big|_d =: LRLR$$

- Two possible linear combinations:

$$\begin{aligned} Q_1(X,Y,Z,W) =& (LRRL - RLLR) + (LRRR - RLLL) \\ &+ (LRLL - RLRR) + (LRLR - RLRL) \end{aligned}$$

$$\begin{aligned} Q_2(X,Y,Z,W) =& (LLLL - RRRR) + (LLLR - RRRL) \\ &+ (LLRL - RRLR) + (LLRR - RRLL) \end{aligned}$$

- After a lengthy calculation:

$$Q_1 = Q_2 = 0$$

uc3m | Universidad Carlos III de Madrid
Departamento de Matemáticas

- Choose a lagrangian

$$\mathfrak{L}(x,v) = \frac{1}{2} g_{ij} v^i v^j + \frac{1}{6} T_{ijk} v^i v^j v^j + \frac{1}{24} C_{ijkl} v^i v^j v^j v^l$$

uc3m | Universidad Carlos III de Madrid
Departamento de Matemáticas

- Choose a lagrangian

$$\mathfrak{L}(x,v) = \frac{1}{2} g_{ij} v^i v^j + \frac{1}{6} T_{ijk} v^i v^j v^j + \frac{1}{24} C_{ijkl} v^i v^j v^j v^l$$

- After a lengthy calculation:

$$Q_1 = Q_2 = 0$$

uc3m | Universidad Carlos III de Madrid
Departamento de Matemáticas

- Choose a lagrangian

$$\mathfrak{L}(x,v) = \frac{1}{2} g_{ij} v^i v^j + \frac{1}{6} T_{ijk} v^i v^j v^j + \frac{1}{24} C_{ijkl} v^i v^j v^j v^l$$

- After a lengthy calculation:

$$Q_1 = Q_2 = 0$$

- What about non-symmetric tensors?

uc3m | Universidad Carlos III de Madrid
Departamento de Matemáticas

References

- F.M. Ciaglia, F. di Cosmo, D. Felice, S. Mancini, G. Marmo and J.M. Pérez-Pardo. *Hamilton-Jacobi approach to Potential Functions in Information Geometry*. **J. Math. Phys. 58,** 063506. (2017).

- P. Facchi, R. Kulkarni, V.I. Man'ko, G. Marmo, E.C.G. Sudarshan and F. Ventriglia. *Classical and quantum Fisher information in the geometrical formulation of quantum mechanics.* **Phys. Lett. A 374,** (2010), 4801-4803