Contraction of Relative Entropy, Riemannian Metrics and Related Measures of Distance between States on Non-Commutative Probability Spaces

Mary Beth Ruskai
University of Massachusetts

Pescara, Italy
June, 2002

References

Monotone metrics introduced by
E.A. Morozova and N.N. Chentshov (1990)

developed by D. Petz and collaborators, e.g.,
"Monotone Metrics on Matrix Spaces"
Lin. Alg. Appl. **244**, 81–96 (1996)
"Geometries of Quantum States"
J. Math. Phys. **37**, 2662–73 (1996).
"Covariance and Fisher Information in Quant Mech." *J. Phys. A* **35**, 929–939 (2002).

A. Lesniewski and M.B. Ruskai, "Relative Entropy and Monotone Riemannian Metrics on Non-Commutative Probability Spaces" *J. Math. Phys.* **40**, 5702–5724 (1999).

M.B. Ruskai, "Contraction of Riemannian Metrics and Related Distance Measures on Pairs of Qubit States" preprint.

A. Jenčová, "Quantum Information Geometry and Standard Purification" *J. Math. Phys.* **43**, 2187–2201 (2002).

States in Quantum Statistical Mechanics
and Quantum Information Theory

Represented by a Density Matrix P
Pos semi-def $P > 0$ and $\text{Tr}\, P = 1$.
Induces map on alg of observables
$A \mapsto \text{Tr}\, PA = f(A)$
Restrict to invertible and fnite-dim
not serious, e.g. $(1-\epsilon)P + \epsilon\frac{1}{d}I$
BUT excludes pure states, i.e., $P^2 = P$.

Want to know how "close" two
(mixed) states P and Q are
See, e.g., Nielsen/Chuang *Quantum Computation ...* Chapter

Expect noise makes states harder
to distinguish, i.e., "distance decreases"

Noise rep by stochastic map Φ
completely positive and trace-preserving

Examples

Relative Entropy

$$H(P,Q) = \mathrm{Tr}\,(P[\log P - \log Q])$$

$H(P,Q) \geq 0$ with $= 0$ iff $P = Q$
BUT not true metric since
  not sym, $H(P,Q) \neq H(Q,P)$ (not serious)
  triangle ineq. not true (even if sym)
Still very useful as a pseudo-distance

Trace norm $\mathrm{Tr}\,|P - Q| \equiv \|P - Q\|_1$
  true metric and gen rel entropy
BUT not smooth; diff contraction props.

Bures metric $2\left[1 - \mathrm{Tr}\,(\sqrt{P}Q\sqrt{P})^{1/2}\right]$
  related to "fidelity" mixed states
Pure $P = |\psi\rangle\langle\psi|, Q = |\Phi\rangle\langle\Phi|$, $F = |\langle\psi, \Phi\rangle|$
true metric – emerges as geodesic dist. of a mono. Riem. metric assoc with a rel entropy

Properties of Relative Entropy

$$P,\ Q > 0 \qquad \mathrm{Tr}\, P = \mathrm{Tr}\, Q$$

a) $H_g(P,Q) \geq \mathrm{Tr}\,(P - Q) = 0$
with $H(P,Q) = 0 \Leftrightarrow P = Q$.
b) $H_g(\lambda P, \lambda Q) = \lambda H(P,Q)$
c) $H_g(P,Q)$ is jointly convex in P and Q
d) monotone: $H_g[\Phi(P), \Phi(Q)] \leq H_g(P,Q)$

In addition we may require
e) symmetric: $H_g(P,Q) = H_g(Q,P)$.
f) differentiable: $f(x,y) = H_g(P+xA, Q+yB)$

Properties (a) – (d) hold for

$H_{\log}(P,Q) = \mathrm{Tr}\,(P[\log P - \log Q])$
corresponds to $g(w) = -\log w$

$H_g(P,Q) = \mathrm{Tr}\,|P-Q|, \qquad g(w) = |w-1|$

convex operator function g such that
$g : (0,\infty) \mapsto \mathbf{R}$, and $g(1) = 0$.

$$H_g(P,Q) = \mathrm{Tr}\,(P^{1/2} g(\Delta_{Q,P}) P^{1/2})$$

mult ops $L_P(A) = PA, \quad R_Q(A) = AQ$

Rel mod op $\Delta_{Q,P} = L_Q R_P^{-1}$

Start with usual Rel Ent

$$H_{\log}(P,Q) = \mathrm{Tr}\,(P[\log P - \log Q])$$

Diff to get Riemannian metric on tang space

$$\begin{aligned} M_p(A,B) &= -\frac{\partial^2}{\partial\alpha\partial\beta} H(P+\alpha A, P+\beta B)\bigg|_{\alpha=\beta=0} \\ &= \mathrm{Tr}\int_0^\infty A\frac{1}{sI+P}B\frac{1}{sI+P}ds \end{aligned}$$

where $\mathrm{Tr}\,A = \mathrm{Tr}\,B = 0, A = A^*, B = B^*$
Extends to bilinear form on full alg s.t

$$M_p(A,B) = \langle A, \Omega_P^g(B)\rangle = \mathrm{Tr}\,A^*\Omega_P^g(B)$$

Defines pos semi-def op (wrt Hilb-Schmidt)

$$\Omega_P^g(B) = \int_0^\infty \frac{1}{sI+P}B\frac{1}{sI+P}ds$$

Note on tang space $M_p(A,B) = M_p(B,A)$
above same as diff $H(P+\alpha B, P+\beta A)$,
$H(P,Q)$ and $H(Q,P)$ give same Riem metric.

leads to geodesic distance, but no simple
closed form expresssion known

$g : (0, \infty) \mapsto \mathbf{R}$ conv. op., $g(1) = 0$.

a) Have pos measure ν_g (or N_g) and a, b, c

$$\begin{aligned} g(w) &= a(w-1) + b(w-1)^2 + c\frac{(w-1)^2}{w} \\ &\quad + \int_0^\infty \frac{(w-1)^2}{w+s} d\nu_g(s) \\ &= a(w-1) + \int_0^\infty \frac{(w-1)^2}{w+s} dN_g(s) \end{aligned}$$

b) A generalized relative entropy $H_g(P, Q)$

$$= \int_0^\infty \text{Tr} \left((Q-P) \frac{1}{L_Q + sR_P} (Q-P) \right) dN_g(s)$$

Note: $g(w) \to wg(w^{-1})$ takes

$$H_g(P, Q) \to H_g(Q, P) = H_{wg(w^{-1})}(P, Q)$$

c) A symmetric generalized relative entropy

$$\begin{aligned} H_g^{\text{sym}}(P, Q) &= H_g(P, Q) + H_g(Q, P) \\ &= H_{g(w)+wg(w^{-1})}(P, Q) \end{aligned}$$

d) fctn $k(w)$ s.t. $wk(w) = k(w^{-1})$, $k(1) = 1$, and $-k(w)$ (or $1/k(w)$) is operator monotone

$$
\begin{aligned}
k(w) &= \frac{g(w) + wg(w^{-1})}{(w-1)^2} \\
&= \int_0^\infty \frac{N_g(s) + s^{-1} N_g(s^{-1})}{s+w} ds \\
&= \int_0^1 \left(\frac{1}{s+w} + \frac{1}{sw+1} \right) \sigma_g(s) ds
\end{aligned}
$$

where $\sigma_g(s) = N_g(s) + s^{-1} N_g(s^{-1})$.

e) family of lin ops Ω_P^g pos semi-def wrt H-S

$$
\begin{aligned}
\Omega_P^g &= \int_0^\infty \left(\frac{1}{sR_P + L_P} + \frac{1}{sL_P + R_P} \right) N_g(s)\, ds \\
&= \int_0^\infty \frac{1}{sR_P + L_P} \sigma_g(s)\, ds
\end{aligned}
$$

f) family of Riemannian metrics dep. on P

$$
\begin{aligned}
M_P^g(A,B) &= -\frac{\partial^2}{\partial\alpha\partial\beta} H_g(P + \alpha A, P + \beta B) \bigg|_{\alpha=\beta=0} \\
&= \langle A, \Omega_P^g(B) \rangle = \mathrm{Tr}\, A^\dagger \Omega_P^g(B)
\end{aligned}
$$

g) A geodesic distance

$$D_g(P,Q) \equiv \text{inf} \int_0^1 \sqrt{\langle \dot{S}(t), \Omega^g_{S(t)} \dot{S}(t) \rangle} dt \quad (1)$$

where the infimum is taken over all smooth paths $S(t)$ with $S(0) = P$ and $S(1) = Q$.

Thm: Relative entropy, Riemannian metrics, and geodesic distances defined above all contract under stochastic maps, i.e,

$$\begin{aligned} H_g[\Phi(P), \Phi(Q)] &\leq H_g(P,Q) \\ M^g_{\Phi(P)}[\Phi(A), \Phi(A)] &\leq M^g_P(A,A) \\ D_g[\Phi(P), \Phi(Q)] &\leq D_g(P,Q) \end{aligned}$$

Thm: Under contract. assump, get 1-1 corr

(c)	$\Leftrightarrow$	(d)	$\Leftrightarrow$	(e) $\Leftrightarrow$	(f)	$\Leftrightarrow$	(g)
H_g^{sym}		$k(w)$		Ω^g_P	$M_P(A,B)$		$D_g(P,Q)$

(d) - (g) due to Petz; (c) to Lesniewski/Ruskai

Defs: For each fixed g and Φ let

$$\eta_g^{\mathsf{RelEnt}}(\Phi) = \sup_{P \neq Q} \frac{H_g[\Phi(P), \Phi(Q)]}{H_g[P, Q]}$$

$$\eta_g^{\mathsf{Riem}}(\Phi) = \sup_{P} \sup_{\substack{A: A = A* \\ \mathrm{Tr}\,(A)=0}} \frac{\langle \Phi(A), \Omega^g_{\Phi(P)}[\Phi(A)] \rangle}{\langle A, \Omega^g_P[A] \rangle}$$

$$\eta_g^{\mathsf{geod}}(\Phi) = \sup_{P \neq Q} \frac{[D_g(\Phi(P), \Phi(Q))]^2}{[D_g(P, Q)]^2}$$

Thm: $\eta_g^{\mathsf{RelEnt}}(\Phi) \geq \eta_g^{\mathsf{Riem}}(\Phi) \geq \eta_g^{\mathsf{geod}}(\Phi)$

Conjecture:

$$\eta_g^{\mathsf{RelEnt}}(\Phi) = \eta_g^{\mathsf{Riem}}(\Phi) = \eta_g^{\mathsf{geod}}(\Phi)$$

Commutative systems, $\langle A, \Omega^g_P[A] \rangle$ Fisher info
$\Rightarrow \langle A, \Omega^g_P[A] \rangle$ indep of g
$\Rightarrow \eta_g^{\mathsf{Riem}}(\Phi)$ indep of g
Thm (Choi, Ruskai, Seneta) $\eta_g^{\mathsf{RelEnt}}(\Phi) = \eta_g^{\mathsf{Riem}}(\Phi)$
$\Rightarrow \eta_g^{\mathsf{RelEnt}}(\Phi)$ indep of g.

$\Phi : \mathcal{A}_1 \mapsto \mathcal{A}_2$ Dual (adjoint) $\widehat{\Phi} : \mathcal{A}_2 \mapsto \mathcal{A}_1$
defined by $\text{Tr}\, A_2^* \Phi(B_1) = \text{Tr}\, \widehat{\Phi}(A_2)^* B_1$

Eigenvalue Equation:

$$[\widehat{\Phi} \circ \Omega^g_{\Phi(P)} \circ \Phi](A) = \lambda \Omega^g_P(A)$$

Easy to check that $A = P$ yields $\lambda = 1$.
Apply max-min

$$\lambda_2(\Phi, P) = \sup_{\substack{A: A = A* \\ \text{Tr}\,(A) = 0}} \frac{\langle \Phi(A), \Omega^g_{\Phi(P)}[\Phi(A)] \rangle}{\langle A, \Omega^g_P[A] \rangle}.$$

Thm: $\eta_g^{\text{Riem}}(\Phi) = \sup_P \lambda_2(\Phi, P) \quad \forall\ g$

Cor: $\eta_{\log}^{\text{Riem}}(\Phi) \leq \sup\limits_{A: \text{Tr}\,(A) = 0} \dfrac{\text{Tr}\, |\Phi(A)|}{\text{Tr}\, |A|} = \eta^{\text{RelEnt}}_{|w-1|}(\Phi)$
$= \frac{1}{2} \sup \{ \text{Tr}\, |\Phi(E - F)| \ : \ \text{1-dim projs},\ EF = 0\}$

The map $A \to (\Omega_P^{\log})^{-1} \circ \widehat{\Phi} \circ \Omega^{\log}_{\Phi(P)} \circ \Phi(A)$
is positivity preserving **only** for $g(w) = -\log w$

Cor. does NOT extend to other g

Rep. of Φ in Pauli basis $\{I, \sigma_x, \sigma_y, \sigma_z\}$

Density matrix $\quad \rho = \frac{1}{2}[I + \mathbf{w} \cdot \sigma]$
ρ a one-dim proj (pure state) $\Leftrightarrow \quad |\mathbf{w}| = 1.$

Gives 1-1 correspondence between states in $\mathbf{C}^2$ and points on unit sphere in $\mathbf{R}^3$

After rotation and diag can assume wlog

$$\mathbf{T} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ t_1 & \lambda_1 & 0 & 0 \\ t_2 & 0 & \lambda_2 & 0 \\ t_3 & 0 & 0 & \lambda_3 \end{pmatrix} = \begin{pmatrix} 1 & \mathbf{0} \\ \mathbf{t} & \Lambda \end{pmatrix}$$

$\mathbf{0}$ and $\mathbf{t}$ are row and column vectors resp.

$$\tfrac{1}{2}\,[I + \mathbf{w}\cdot\sigma] \quad \xrightarrow{\Phi} \quad \tfrac{1}{2}\left[I + \sum_k (t_k + \lambda_k w_k)\sigma_k\right]$$

Φ maps pure states to ellipsoid in Bloch sphere

$$\left(\frac{x_1 - t_1}{\lambda_1}\right)^2 + \left(\frac{x_2 - t_2}{\lambda_2}\right)^2 + \left(\frac{x_3 - t_3}{\lambda_3}\right)^2 = 1$$

BUT not every ellipsoid from Comp. Pos. Φ
actual conditions complicated

Example 1: Unital qubit channel

B real 3×3 matrix with $\|B\| \leq 1$

$$\Phi : \tfrac{1}{2}[I + \mathbf{w} \cdot \sigma] \mapsto \tfrac{1}{2}[I + (B\mathbf{w}) \cdot \sigma]$$

(comp. pos. not needed) $\Phi(I) = I$

$$\eta_g^{\text{RelEnt}}(\Phi_B) = \eta_g^{\text{Riem}}(\Phi_B) = \eta_g^{\text{geod}}(\Phi_B) = \|B\|^2$$

$$\begin{aligned} \eta_g^{Riem}(\Phi_B) &= \|B\|^2 = \text{largest e-val } B^*B \quad \forall g \\ &= \text{2nd largest e-val} \begin{pmatrix} 1 & 0 \\ 0 & B^* \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & B \end{pmatrix} \end{aligned}$$

$$\begin{aligned} \eta_g^{\text{Riem}}(\Phi) &\leq \eta\text{Dob}(\Phi) = \|B\| \\ &\equiv \sup_{A:\text{Tr}\,(A)=0} \frac{\text{Tr}\,|\Phi(A)|}{\text{Tr}\,|A|} \end{aligned}$$

Consistent with $\quad \|B\| \leq 1 \quad \Rightarrow \quad \|B\|^2 \leq \|B\|$

Example 2: Commutative 2-dim : or

QC: $\Phi : \ \frac{1}{2}[I + \mathbf{w}\cdot\sigma] \ \mapsto \ \frac{1}{2}[I + (t_3 + \lambda_3 w_3)\sigma_3]$

QC means Quantum $\to$ Classical
CC means Classical $\to$ Classical

$$\eta_g(\Phi_{QC}) = \eta_g(\Phi_{CC}) \quad \text{since} \quad \text{e.g.,}$$

$$H_g(P_{\mathbf{w}}, P_{\mathbf{x}}) \geq H_g(\tfrac{1}{2}[I + w_3\sigma_3], \tfrac{1}{2}[I + x_3\sigma_3])$$

$$\frac{1}{2}[I + w\sigma_3] = \frac{1}{2}\begin{pmatrix} 1+w & 0 \\ 0 & 1-w \end{pmatrix} \ \leftrightarrow \ \frac{1}{2}\begin{pmatrix} 1+w \\ 1-w \end{pmatrix}$$

map above corresponds to column stochastic

$$\frac{1}{2}\begin{pmatrix} 1+\lambda+t & 1-\lambda+t \\ 1-\lambda-t & 1+\lambda-t \end{pmatrix}\frac{1}{2}\begin{pmatrix} 1+w \\ 1-w \end{pmatrix} = \frac{1}{2}\begin{pmatrix} 1+t+\lambda w \\ 1-t-\lambda w \end{pmatrix}$$

$$\begin{aligned} \eta_g^{\text{RelEnt}}(\Phi_A) &= \eta_g^{\text{Riem}}(\Phi_A) \\ &= \frac{2\lambda^2}{\left[\sqrt{(1+\lambda)^2 - t^2} + \sqrt{(1-\lambda)^2 - t^2}\right]^2} \end{aligned}$$

reduces to λ^2 when $t = 0$ (recover unital).

Example 3: Non-unital qubit CQ

CQ means Classical $\rightarrow$ Quantum

$$\Phi : \ \tfrac{1}{2}[I + \mathbf{w}\cdot\sigma] \ \mapsto \ \tfrac{1}{2}[I + t_1\sigma_1 + \lambda_3 w_3\, \sigma_3]$$

$$= \ \tfrac{1}{2}\begin{pmatrix} 1+\lambda_3 w_3 & t_1 \\ t_1 & 1-\lambda_3 w_3 \end{pmatrix}$$

Comp. Pos. iff $\lambda_3^2 + t_1^2 \leq 1$ for CQ

compare $|\lambda_3| + |t_3| \leq 1$ for QC.

$$\eta^{\text{Riem}}_{(w-1)^2}(\phi) = \frac{\lambda^2}{1-t^2} \qquad = 1 \ \text{ if } \lambda_3^2 + t_1^2 = 1$$

Can find g such that $\eta^{\text{Riem}}_{(w-1)^2}(\phi) > \eta^{\text{Riem}}_g(\phi)$

i.e. • $\eta^{\text{Riem}}_g(\phi)$ depends non-triv on g.

$\eta^{\text{Dob}}(\phi) = \lambda$ but not general upper bound

$\eta^{\text{Dob}}(\phi) < \eta^{\text{Riem}}_{(w-1)^2}(\phi)$ when $\lambda^2 \leq 1 - t^2 < \lambda$