Non-commutative optimization & Scaling

G group, $\pi: G \to GL(V)$ representation, $v \in V$ vector

$$\boxed{\mathrm{Cap}(v) = \inf_{g \in G} \log \|\pi(g) v\|}$$

e.g. operator scaling: $G = SL(n) \times SL(n)$ $\quad V = \mathrm{Mat}_n(\mathbb{C})^m$

$$\pi(L,R)(A_1,\dots,A_m) = (LA_1R^{-1},\dots,LA_mR^{-1})$$

Motivation:

* marginal + scaling problems: (Rafael)

  $\mathrm{Cap}(v) > -\infty \iff$ can scale v to "d.s."

  $\mathrm{Cap}(v) = \log\|v\| \iff v$ is "doubly stochastic"

  Cap = natural potential fn

* invariant theory: (Ankit)

  $\mathrm{Cap}(v) = -\infty \iff v \in$ null cone

  and moment polytopes

* noncommutative geometric programming!

This talk: WHY + TOOLS: CONVEXITY! $\to$ continuous ALGOS

Commutative groups: e.g. $G = T(n) = \left\{ z = \begin{pmatrix} z_1 & & \\ & \ddots & \\ & & z_n \end{pmatrix} \right\}$

$$v = \sum_{\omega \in \Omega} v_\omega X^\omega \quad \text{Laurent polynomial } (\Omega \subseteq \mathbb{Z}^n)$$

$$\pi(z) v = \sum_\omega (v_\omega z^\omega) X^\omega \quad \leftarrow \text{rescale variables}$$

Any G-repr. is "of this form"?!

$$\mathrm{Cap}(v) = \inf_z f(z), \quad f(z) = \frac{1}{2} \log \overbrace{\sum_\omega |v_\omega|^2 |z_1|^{2\omega_1} \cdots |z_n|^{2\omega_n}}^{\text{posynomial in } |z_i|^2 > 0}$$

↳ (unconstrained) geometric programming!

change of var's: $z = e^h$, $h \in \mathbb{R}^n$ WLOG

$$f(e^h) = \frac{1}{2} \log \sum_\omega |v_\omega|^2 e^{2h\cdot\omega}$$

↳ Convex in h ⁈ ⇝ efficient algos (Nisheeth, Gurvits)

* $\mathrm{Cap}(v) > -\infty$ iff $0 \in \Delta(v) := \mathrm{conv}\{\omega : v_\omega \neq 0\}$

Farkas' Newton polytope of v

coeff's of scaled poly

Gradient (w.r.t. h): $\nabla f(z) = \sum_\omega \omega \frac{|v_\omega|^2 |z^\omega|^2}{\sum_\omega |v_\omega|^2 |z^\omega|^2} \in \Delta(v)$

* can approx. scale to ANY point in $\Delta(v)$ i.e. $\forall \varepsilon, p \in \Delta(v)$ $\exists z : \nabla f(z) \approx_\varepsilon p$

* e.g., for matrix scaling

$$G = \{T(n) \times T(n) \ni (L,R) : \det = 1\} \quad V = \mathrm{Mat}_n(\mathbb{C}) \ni A$$

$$\pi(L,R) = LAR^{-1} = B$$

⇝ $\nabla f(L,R)$ = traceless part of row&col sums of $\frac{|B_{ij}|^2}{\|B\|_F^2}$ $\in \mathbb{R}^n \times \mathbb{R}^n$

gradients are interesting!

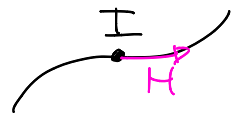
<u>Noncommutative groups</u>: e.g. GL_n, SL_n, products, ...

$$\mathrm{cap}(v) = \inf_g f(g), \quad f(g) = \frac{1}{2}\log \|\pi(g)\cdot v\|^2$$

Change of vars? $GL_n = \{ g = u e^H \mid u \text{ unitary}, H = H^* \}$

$$f(e^H) = \frac{1}{2}\log\langle v, \pi(e^{2H}) v\rangle$$ ↯ NOT convex

HOWEVER: $\boxed{f(e^{Ht})}$ is <u>convex</u> in t, i.e. f is convex along <u>geodesics</u> $\{e^{Ht}\} \subseteq GL_n$!



Pf: WLOG $\|v\|=1$ & $t=0$.

$$f(e^{Ht}) = \frac{1}{2}\log\langle v, \pi(e^{2Ht})v\rangle = \frac{1}{2}\underbrace{\langle v, e^{2\tilde{H}t} v\rangle}_{\approx \text{cumulant gen. fn.}}$$

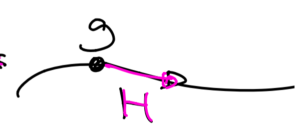
$$\partial_t f(e^{Ht}) = \frac{\langle v, \tilde{H} e^{2\tilde{H}t} \cdot v\rangle}{\langle v, e^{2\tilde{H}t}\cdot v\rangle}$$

$$\partial^2_{t=0} f(e^{Ht}) = \ldots = 2\langle v, \underbrace{\overbrace{(\tilde{H} - \langle v, \tilde{H} v\rangle)^2}^{\text{variance}}}_{\geq 0} v\rangle \geq 0 \qquad \square$$

↳ <u>g-convex optimization</u> → grad flow always converges
→ Rafael, WIP, ...

<u>Gradient</u> $\nabla f(g)$? $n\times n$ Herm. matrix (for $G = GL_n$) s.th.

$$\mathrm{tr}[\nabla f(g) H] = \partial_{t=0} f(e^{Ht} g) \quad \forall H = H^*$$



* "moment map", often has interpretation as "marginal"

* e.g., for <u>operator scaling</u>:

$G = SL(n)\times SL(n)$ $\quad V = \mathrm{Mat}_n^m(\mathbb{C}) \ni A = (A_1,\dots,A_m)$

$\pi(L,R)\cdot A = (LA_1R^{-1},\dots,LA_mR^{-1}) = B$

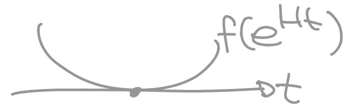
$\rightsquigarrow \nabla f(L,R) =$ traceless part of $\left(\sum_i B_iB_i^*, -\sum_i B_i^*B_i\right)\frac{1}{\|B\|^2}$

grad's are interesting

In general: $\boxed{\nabla f(g) = 0 \iff \mathrm{cap}(v) = f(g)}$

<u>Pf</u>: ($\Rightarrow$) WLOG $g = I$. For all $g' = ue^H$:

$f(g') = f(\cancel{u}e^H) \geq f(I)$



$\square$

Similar: In $\overline{\pi(GL_n)\cdot v}$, min. norm vectors form single U_n-orbit.

<u>Noncommutative duality</u>: Two dual opt. problems:

① $\mathrm{cap}(v) = \inf_g f(g) > -\infty$

② $\mathrm{ds}(v) = \inf_g \|\nabla f(g)\|_F \overset{!}{=} 0$

minimize norm

NC Farkas (Kempf-Ness)

Scale to d.s.

* $f(g) - \mathrm{cap}(v)$ vs. $\|\nabla f(g)\|_F - \mathrm{ds}(v)$

* $\exists$ gaps Γ, Γ': $\begin{cases} \|\nabla f(g)\| \leq \Gamma \Rightarrow \mathrm{ds}(v) = 0 \\ v \text{ integral}, \ f(g) \leq \Gamma' \Rightarrow \mathrm{cap}(v) = -\infty \end{cases}$

↳ decision problem in NP ∩ coNP? bit complexity..

Scaling to other marginals? $\{\nabla f(g)\} = ?$

* $\Delta(V) = \overline{\{\operatorname{spec} \nabla f(g) \mid g \in G\}} \subseteq \mathbb{R}^n$ polytope!

  = all marginals we can scale to (q. Newton poly)

  "moment polytope" (Mumford, Kirwan)

* e.g. tensors, Horn, quivers, Brascamp–Lieb, ... Matthias, Nisheeth, Cole
* can have exp. many vertices, faces (in dim V)

* ALL points characterized by g-convex optimization

$$x \in \Delta(V) \iff \operatorname{cap}_x(V) = \inf_{g \text{ lower tri}} \left(\prod g_{ii}^{-2x_i} \right) \|\pi(g)v\|^2 > 0$$

OPEN Q: Fast algos? Other opti problems w/ symmetry?

NO TIME FOR: INVARIANTS