

Hadamard 鏡像降下法による テンソルパラメタの計算

坂部圭哉 (ミュンヘン大学)

共同研究者：M. Levent Doğan & Michael Walter

数理最適化で繋ぐ幾何解析と応用数学@名古屋大学

2026 年 9 月 11 日



arXiv:2609.06633

- ① 鏡像降下法 (Euclid の場合)
- ② 双対変数による定式化
- ③ Hadamard 多様体への拡張
- ④ 応用：エンタングルメント多面体上の凸最適化

制約つき凸最適化問題：

$$\text{minimize: } Q(p) \quad \text{s.t. } p \in C$$

- 制約集合 $\emptyset \neq C \subseteq \mathbb{R}^n$: 凸・閉

- 目的関数 $Q: \mathbb{R}^n \rightarrow (-\infty, +\infty]$:

- 凸・下半連続
- $C \subseteq \text{dom } Q := \{p \in \mathbb{R}^n \mid Q(p) < \infty\}$
- $\text{int dom } Q$ 上で可微分: $dQ(p) \in (\mathbb{R}^n)^*$

微分は双対空間の元

$$\text{ここで, } \underbrace{\langle dQ(p), q \rangle_{(\mathbb{R}^n)^*, \mathbb{R}^n}}_{\text{主双対積}} := (dQ(p))[q] = \underbrace{(\nabla Q(p), q)_{\mathbb{R}^n}}_{\text{ユークリッド内積}}$$

主双対積

ユークリッド内積

課題：制約集合 C をどう扱うか？

Bregman ダイバージェンス

3/34

アイデア：制約集合 C に **Bregman ダイバージェンス** (距離のようなもの) を導入
 $\implies C$ 上に “自然な幾何構造” を与える

Bregman ダイバージェンス

$h: \mathbb{R}^n \rightarrow (-\infty, +\infty]$ (凸, $\text{int dom } h$ 上で可微分) に対し

$$D_h(p\|q) := h(p) - h(q) - \langle dh(q), p - q \rangle, \quad (p \in \mathbb{R}^n, q \in \text{int dom } h).$$

($p \notin \text{dom } h$ なら $D_h(p\|q) := \infty$ とする)



※ 「二乗距離」の一般化： $h(p) = \|p\|^2/2$ ならば, $D_h(p\|q) = \|p - q\|^2/2$.

C 上の自然な距離尺度とするため, h は以下を満たすようにする：

- $\overline{\text{dom } h} = C$
 - 狭義凸
 - $\|dh(p_i)\| \rightarrow \infty$ as $p_i \rightarrow \partial C$ (本質的平滑)
- } $dh: \text{int dom } h \rightarrow (\mathbb{R}^n)^*$ が全単射

再掲： minimize: $Q(p)$ s.t. $p \in C$

鏡像降下法 / Bregman 近接勾配法

$\eta > 0$: ステップ幅

$$p_{k+1} := \operatorname{argmin}_{p \in \mathbb{R}^n} \left\{ \overbrace{Q(p_k) + \langle dQ(p_k), p - p_k \rangle}^{Q \text{ の一次近似}} + \overbrace{\frac{1}{\eta} D_h(p \| p_k)}^{\text{近接項}} \right\} \quad (1)$$

一次最適性条件を取ることで、「双対空間での更新」と同値：

$$dh(p_{k+1}) = dh(p_k) - \eta dQ(p_k) \quad (2)$$

※ (1) の定式化が「Bregman 近接勾配法」、(2) が「鏡像降下法」と呼ばれやすい。

※歴史的には、(2) が [Nemirovsky and Yudin, 1983] により提案された後、同値性が [Beck and Teboulle, 2003] で示された。

ステップ幅 $\eta > 0$ を選ぶ基準として、**相対的平滑性**が有用：

相対的平滑性 [Bauschke et al., 2017, Lu et al., 2018]

目的関数 Q が h に対して **相対的 L -平滑**

$$\begin{aligned} &\stackrel{\text{def.}}{\iff} D_Q(p\|q) \leq L \cdot D_h(p\|q) \quad (\forall p \in \mathbb{R}^n, q \in \text{int dom } h) \\ &\iff Lh - Q \text{ が凸} \end{aligned}$$

- $h(p) = \|p\|^2/2$ のとき、 L -平滑性と同値： $D_Q(p\|q) \leq L\|p - q\|^2/2$.
- 一般に、 Q が L -平滑、 h が μ -強凸ならば、相対的 L/μ -平滑：

$$D_Q(p\|q) \leq L\|p - q\|^2/2 \leq (L/\mu)D_h(p\|q).$$

☆ 相対的 L -平滑 $\implies \eta \leq 1/L$ と取ると良い (次頁へ)

鏡像降下法の収束性 [Bauschke et al., 2017, Lu et al., 2018]

(Q, h) : 相対的 L -平滑 $\implies \eta \leq 1/L$ と取る.
すべての $q \in \mathbb{R}^n$ に対し,

$$Q(p_k) - Q(q) \leq \frac{D_h(q \| p_0)}{\eta k}.$$

※非自明な主張になるのは $D_h(q \| p_0) < \infty$ の場合 ($\approx q \in C$ の場合) のみ

- ① 鏡像降下法 (Euclid の場合)
- ② 双対変数による定式化
- ③ Hadamard 多様体への拡張
- ④ 応用：エンタングルメント多面体上の凸最適化

問題： minimize: $Q(p)$ s.t. $p \in C \subseteq \mathbb{R}^n$

距離生成関数 h ： $\overline{\text{dom } h} = C$, $\text{d}h: \text{int dom } h \rightarrow (\mathbb{R}^n)^*$ が全単射

相対的 L -平滑： $D_Q(p\|q) \leq L \cdot D_h(p\|q)$

更新式： $\text{d}h(p_{k+1}) = \text{d}h(p_k) - \eta \text{d}Q(p_k) \quad (\eta \leq 1/L)$

収束保証： $Q(p_k) - Q(q) \leq \frac{D_h(q\|p_0)}{\eta k} \quad (\forall q \in C)$

h の代わりに Legendre–Fenchel 共役 $h^* =: f$ を用いて書き直す

- 問題が無制約に
- 更新式が陽的に
- 多様体への (我々の方法の) 一般化に向けて

Legendre–Fenchel 共役

$F: \mathbb{R}^n \rightarrow (-\infty, +\infty]$, 下半連続, 凸, $\text{dom } F \neq \emptyset$

Legendre–Fenchel 共役 $F^*: (\mathbb{R}^n)^* \rightarrow (-\infty, +\infty]$ を以下で定義:

$$F^*(x) := \sup_{p \in \mathbb{R}^n} \langle x, p \rangle - F(p)$$

性質:

- $(F^*)^* = F$

微分可能性のもと,

- $\text{d}F^* = (\text{d}F)^{-1}$
- $D_F(\text{d}F^*(x) \| \text{d}F^*(y)) = D_{F^*}(y \| x)$
- $\text{dom } F = \mathbb{R}^n$ ならば, $\overline{\text{dom } F^*} = \overline{\text{d}F(\mathbb{R}^n)} := \overline{\{\text{d}F(p) \mid p \in \mathbb{R}^n\}}$ (微分の集合)
(定義から $\text{dom } F^* \supseteq \text{d}F(\mathbb{R}^n)$. 逆はやや非自明 (e.g., [Rockafellar, 1970, § 24]))

書き換えたい： $Q(p_k) - Q(q) \leq \frac{D_h(q\|p_0)}{\eta k} \quad (\forall p \in C)$

素直には、 $Q(df(x_k)) - Q(df(y)) \leq D_f(x_0\|y)/(\eta k) \quad (\forall y \in (\mathbb{R}^n)^*)$.

これは正しいが、最適解 y^* が存在しない可能性によりあまり実用的ではない。

代わりに、任意の $q \in C$ に対してはじめの式を直接書き換える：

$$\begin{aligned} D_h(q\|p_0) &= -h(p_0) - \langle dh(p_0), q - p_0 \rangle + h(q) \\ &= f(x_0) - \langle x_0, p_0 \rangle - \langle x_0, q - p_0 \rangle + \sup_{x \in (\mathbb{R}^n)^*} (\langle x, q \rangle - f(x)) \\ &= (f(x_0) - \langle x_0, q \rangle) - \inf_{x \in (\mathbb{R}^n)^*} (f(x) - \langle x, q \rangle) \end{aligned}$$

よって以下を得る：

$$Q(df(x_k)) - Q(q) \leq \frac{f^q(x_0) - \inf_{x \in (\mathbb{R}^n)^*} f^q(x)}{\eta k} \quad (f^q := f - \langle \cdot, q \rangle)$$

簡単のため, $E := (\mathbb{R}^n)^*$ と書く

問題設定:

$$\text{minimize: } Q(df(x)) \text{ s.t. } x \in E$$

- $f: E \rightarrow \mathbb{R}$: 凸, 可微分
- $Q: E^* \rightarrow (-\infty, +\infty]$: 下半連続, 凸, $\text{int dom } Q$ 上で可微分
- $df(E) \subseteq \text{int dom } Q$

鏡像降下法 (双対空間で): $x_{k+1} := x_k - \eta dfQ(df(x_k))$

相対的 L -平滑: $D_Q(df(y) \| df(x)) \leq L \cdot D_f(x \| y)$

収束性: $\eta \leq 1/L$ のもと, $\forall q \in E^*$ に対し,

$$Q(df(x_k)) - Q(q) \leq \frac{f^q(x_0) - \inf_{x \in E} f^q(x)}{\eta k} \quad (f^q := f - \langle \cdot, q \rangle)$$

※ x_k そのものは一般には収束しない

例：ノルム最小化

12/34

例： $Q(p) = \|p\|^2/2$

問題：

$$\text{minimize}_{x \in E} \quad Q(df(x)) = \|\nabla f(x)\|^2/2$$

仮定： f : L -平滑

$\implies (Q, f)$: 相対的 L -平滑

$\because$ 一般に、 f が L_f -平滑 & Q が L_Q -平滑 $\implies$ 相対的 $(L_f L_Q)$ -平滑：

$$D_Q(df(y) \| df(x)) \leq (L_Q/2) \|df(x) - df(y)\|^2 \leq L_f L_Q D_f(x \| y)$$

更新式： $x_{k+1} := x_k - \eta df(x_k) = x_k - \eta \nabla f(x_k)$ f に対する勾配降下法！

収束性： $\eta \leq 1/L$ ならば、

$$\frac{\|\nabla f(x_k)\|^2}{2} \leq \frac{f(x_0) - \inf_{x \in E} f(x)}{\eta k}$$

「勾配降下法は、凸関数の勾配を最小化する」が再現された！！

(cf. $\inf f = -\infty$ の場合：[Hirai and Sakabe, 2024])

- ① 鏡像降下法 (Euclid の場合)
- ② 双対変数による定式化
- ③ Hadamard 多様体への拡張
- ④ 応用：エンタングルメント多面体上の凸最適化

前節の問題設定：

$$\text{minimize: } Q(df(x)) \text{ s.t. } x \in E$$

多様体 (特に, Hadamard 多様体) $\mathcal{M}$ に拡張したい

- 問題や更新式はどのように一般化される？
- f は Q はどのような関数か？
- Bregman ダイバージェンス・相対的平滑性などの条件はどうなる？

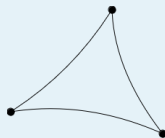
ここでは, [Hirai, 2025] の方法で問題を一般化し,
本研究 [Doğan, Sakabe, and Walter, 2026] で離散時間アルゴリズムを与える.

※通常の意味 (?) の「鏡像降下法が多様体への一般化」(e.g., [Jiang et al., 2026])
とは異なる. (主空間の一般化 vs 双対空間の一般化)

Hadamard 多様体

以下を満たす Riemann 多様体 $\mathcal{M}$

- 完備
 - 単連結
 - 断面曲率 ≤ 0
- } CAT(0): 任意の三角形が「潰れている」



☆任意の二点 $x, y \in \mathcal{M}$ をつなぐ測地線が唯一存在

例: $\mathbb{R}^n$: Euclid 空間, $\mathbb{H}^n$: 双曲空間,
PD(n): $n \times n$ 正定値エルミート行列全体の集合

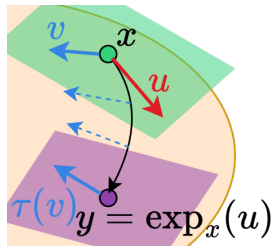
記号：

- $T_x \mathcal{M}$: 接空間 ($x \in \mathcal{M}$)
- $T\mathcal{M} := \sqcup_{x \in \mathcal{M}} T_x \mathcal{M}$: 接バンドル
- $T_x^* \mathcal{M} := (T_x \mathcal{M})^*$: 余接空間
- $T^* \mathcal{M}$: 余接バンドル

指数写像

$$\exp_x: T_x \mathcal{M} \rightarrow \mathcal{M}$$

$\exp_x(u) := u/\|u\|$ 方向の測地線に沿って距離 $\|u\|$ 進んだ点



平行移動

- $\tau_{x \rightarrow y}: T_x \mathcal{M} \rightarrow T_y \mathcal{M}$: 測地線に沿った接ベクトルの平行移動
- $\tau_{x \rightarrow y}^*: T_y^* \mathcal{M} \rightarrow T_x^* \mathcal{M}$: 上の随伴写像：余接ベクトルの平行移動

[Hirai, 2025] の提案「 Q -ノルム最小化問題」:

$$\text{minimize: } Q(df(x)) \quad \text{s.t. } x \in \mathcal{M} : \text{Hadamard 多様体}$$

ここで,

- $f: \mathcal{M} \rightarrow \mathbb{R}$: 測地的凸 (任意の測地線に沿って凸), 可微分
- "norm-like function" $Q: T^*\mathcal{M} \rightarrow (-\infty, +\infty]$ は以下を満たす:¹
 - 各点 $x \in \mathcal{M}$ で, $Q_x: T_x^*\mathcal{M} \rightarrow (-\infty, +\infty]$: 下半連続, 凸, $\text{int dom } Q_x$ で可微分
 - 平行移動に対して不変: $Q_x = Q_y \circ \tau_{y \rightsquigarrow x}^*$

任意の曲線に沿った平行移動

後者の条件から, ある $x_0 \in \mathcal{M}$ での Q_{x_0} から Q 全体が定まる.

上の条件を満たす Q 定めるための Q_{x_0} の必要十分条件は,

- $Q_{x_0}: T_{x_0}^*\mathcal{M} \rightarrow (-\infty, +\infty]$: 下半連続, 凸, $\text{int dom } Q_{x_0}$ で微分可能
- ホロノミー群に対して不変: $Q_{x_0} \circ \tau_{x_0 \rightsquigarrow x_0}^* = Q_{x_0}$ (cf. $\text{Hol}_x^*(\mathcal{M}) := \{\tau_{x_0 \rightsquigarrow x_0}^*\}$)
- $df(x) \in \text{int dom } Q_x$ ($\forall x \in \mathcal{M}$)

¹ Q の正則性や非負性などに元論文 [Hirai, 2025] と差異あり.

再掲：

$$\text{minimize: } Q(df(x)) \quad \text{s.t. } x \in \mathcal{M}$$

[Hirai, 2025] はこの最小化問題に対し、 Q -勾配流² を提案：

Q -勾配流 (Q -gradient flow)

$$\dot{x}_t = -dQ_{x_t}(df(x_t)) \quad (\in T_{df(x)}^* T_x^* \mathcal{M} \cong (T_x^* \mathcal{M})^* \cong T_x \mathcal{M})$$

命題 [Hirai, 2025]

- $Q(df(x_t))$ は単調非増加
- 広いクラスの Hadamard 多様体 ($\mathbb{R}^n, \mathbb{H}^n, \text{PD}(n)$ やその直積を含む) で,
 $Q(df(x_t)) \rightarrow \inf_{x \in \mathcal{M}} Q(df(x)).$

※鏡像降下法の連続極限との関係は指摘されていなかった。

²元論文 [Hirai, 2025] では、 Q の代わりに $Q^2/2$ を使用．元論文ではさらに $Q \geq 0$ を仮定していたため、 Q の最小化と Q^2 の最小化は等価だった．

本研究 [Doğan, Sakabe, and Walter, 2026] では,

- Q -勾配流が, 鏡像降下法連続時間極限の一般化とみなせることを指摘.
- Euclid の鏡像流と同様の形で, **定量的な収束保証をすべての Hadamard 多様体**で示す.
- 鏡像降下法に対応する自然な離散化を提案し, 収束保証を与える.

"Hadamard 鏡像降下法 (Hadamard mirror descent)"

再掲 **指数写像に一般化** 下 **Riemannian Bregman ダイバージェンス (次頁) に一般化**

更新式: $x_{k+1} := x_k - \eta \mathrm{d}Q(\mathrm{d}f(x_k))$

相対的 L -平滑: $D_Q(\mathrm{d}f(y) \| \mathrm{d}f(x)) \leq L \cdot D_f(x \| y)$

収束性: $\eta \leq 1/L$ のもと, $\forall q \in E^*$ に対し **Busemann 関数 (次頁) を用いて定義**

$$Q(\mathrm{d}f(x_k)) - Q(q) \leq \frac{f^q(x_0) - \inf_{x \in E} f^q(x)}{\eta k} \quad (f^q := f - \langle \cdot, q \rangle)$$

定義

- $f: \mathcal{M} \rightarrow \mathbb{R}$ 測地的凸, 可微分

Riemannian Bregman ダイバージェンス:

$$D_f^{\mathcal{M}}(x \| y) := f(x) - f(y) - \langle df(y), \exp_y^{-1}(x) \rangle$$

- これを用いて, (Q, f) が**相対的 L -平滑**

$$\stackrel{\text{def.}}{\iff} D_{Q_x}(\tau_{x \rightarrow y}^* df(y) \| df(x)) \leq L \cdot D_f^{\mathcal{M}}(x \| y) \quad (\forall x, y \in \mathcal{M})$$

相対的 L -平滑性の十分条件のために, 以下の概念が有用:

- $f: \mathcal{M} \rightarrow \mathbb{R}$ が **L -双対平滑 (dual-smooth)**

$$\stackrel{\text{def.}}{\iff} D_f^{\mathcal{M}}(x \| y) \geq (1/2L) \|\tau_{x \rightarrow y}^* df(y) - df(x)\|^2 \quad (\forall x, y \in \mathcal{M})$$

- $f: L_f$ -双対平滑 & $Q: L_Q$ -平滑 $\implies (Q, f)$: 相対的 $(L_f L_Q)$ -平滑:
 $(D_{Q_x}(\tau_{x \rightarrow y}^* df(y) \| df(x)) \leq (L_Q/2) \|\tau_{x \rightarrow y}^* df(y) - df(x)\|^2 \leq (L_f L_Q) D_f^{\mathcal{M}}(y \| x))$
- $\mathcal{M} \cong \mathbb{R}^n$ のとき, L -双対平滑 $\iff L$ -平滑
- 一般の Hadamard 多様体 $\mathcal{M}$ では, L -双対平滑 $\implies L$ -平滑 のみ

線形 (or アフィン) 関数の Hadamard 多様体への一般化

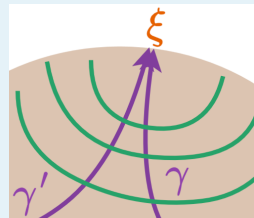
Busemann 関数

$\gamma: \mathbb{R}_{\geq 0} \rightarrow \mathcal{M}$ 測地的半直線 に対し,
Busemann 関数 $b^\gamma: \mathcal{M} \rightarrow \mathbb{R}$ を以下で定義:

$$b^\gamma(x) := \lim_{t \rightarrow \infty} (\|\dot{\gamma}\| \text{dist}(x, \gamma(t)) - \|\dot{\gamma}\|^2 t).$$

γ, γ' が **漸近的** ($\limsup_{t \rightarrow \infty} \text{dist}(\gamma(t), \gamma'(t)) < \infty$) $\implies b^\gamma - b^{\gamma'} = \text{const.}$

ゆえに, 大域定数を除いて, 漸近関係からなる**同値類全体の集合**
 CM^∞ でパラメトライズされる: $CM^\infty \ni \xi \mapsto b^\xi$.



- 測地的凸
- $b^\xi(x)$ は x が ξ の"方向"からどれほど遠いかを表す
- $\mathcal{M} = \mathbb{R}^n$ のとき, CM^∞ は測地線の"速度" ξ で特徴づけられる.
 $\implies b^\xi(\cdot) = -\langle \xi, \cdot \rangle + \text{const.}$

"norm-like function" Q について、以下が示せる：

補題

Busemann 関数 b^ξ について、 $Q(-db^\xi(x)) = Q(-db^\xi(y))$ ($\forall x, y \in \mathcal{M}$). ゆえに、

$$Q(\xi) := Q(-db^\xi(x))$$

が well-defined.

※負号は、 $\mathcal{M} \cong \mathbb{R}^n$ のとき $b^\xi(\cdot) = -\langle \xi, \cdot \rangle + \text{const.}$ と整合.

収束レート (本研究)

ステップ幅 $\eta \leq 1/L$ と取るとき、任意の Busemann 関数 b^ξ について

$$Q(df(x_k)) - Q(\xi) \leq \frac{f^\xi(x_0) - \inf_{x \in \mathcal{M}} f^\xi(x)}{\eta k} \quad (f^\xi := f + b^\xi)$$

特に、 $Q(df(x_k)) \rightarrow \inf_{x \in \mathcal{M}} Q(df(x))$

問題設定：

$$\text{minimize: } Q(df(x)) \text{ s.t. } x \in \mathcal{M}$$

- $f: \mathcal{M} \rightarrow \mathbb{R}$: 測地的凸, 可微分
- $Q: T^*\mathcal{M} \rightarrow (-\infty, +\infty]$
 - それぞれの $Q_x: T_x^*\mathcal{M} \rightarrow (-\infty, +\infty]$: 下半連続, 凸, $\text{int dom } Q_x$ 上で可微分
 - $\mathcal{M}$ 上の平行移動に対して不変
- $df(x) \subseteq \text{int dom } Q_x$

Hadamard 鏡像降下法： $x_{k+1} := \exp_{x_k}(-\eta dQ_{x_k}(df(x_k)))$

相対的 L -平滑： $D_{Q_x}(\tau_{x \rightarrow y}^* df(y) \| df(x)) \leq L \cdot D_f^{\mathcal{M}}(x \| y)$

収束性： $\eta \leq 1/L$ のもと, $\forall \xi \in C\mathcal{M}^\infty$ に対し,

$$Q(df(x_k)) - Q(\xi) \leq \frac{f^\xi(x_0) - \inf_{x \in \mathcal{M}} f^\xi(x)}{\eta k} \quad (f^\xi := f + b^\xi)$$

※連続時間版も成立 (省略)

- ① 鏡像降下法 (Euclid の場合)
- ② 双対変数による定式化
- ③ Hadamard 多様体への拡張
- ④ 応用：エンタングルメント多面体上の凸最適化

$\mathrm{PD}(n)$: $n \times n$ 正定値エルミート行列全体の集合

これは、以下の計量のもと Hadamard 多様体になる：

$$(X, Y)_{T_x \mathrm{PD}(n)} := \mathrm{tr}[x^{-1} X x^{-1} Y] \quad (X, Y \in T_x \mathrm{PD}(n) \cong \mathrm{Herm}(n))$$

- 代表的な**対称空間**の例
- 任意の $g \in \mathrm{GL}(n)$ に対し, $x \mapsto g x g^\dagger$ は等長写像
- 測地線は $t \mapsto g e^{\mathrm{diag} z} g^\dagger$ で与えられる ($g \in \mathrm{GL}(n), z \in \mathbb{R}^n$).

$\mathcal{M} = \text{PD}(n)$ の場合の "norm-like function" Q はどのように特徴づけられるか？

事実^a : $I \in \text{PD}(n)$ でのホロノミ一群は、ユニタリ共役変換 :

$$\text{Hol}_I^*(\text{PD}(n)) = \{X \mapsto u^\dagger X u \mid u \in \text{U}(n)\}$$

(ここで, $T_I^* \text{PD}(n) \underset{\text{計量}}{\cong} T_I \text{PD}(n) \cong \text{Herm}(n)$ と同一視)

^a例えば, [Sakai, 1996, Thm. IV.6.9]

Q は, ユニタリ共役不変な凸関数 $Q_I = \Phi: \text{Herm}(n) \rightarrow (-\infty, +\infty]$ から誘導される.

Φ はユニタリ共役不変のため, 入力の特クトル (固有値集合) の関数 $\implies$

Φ は**対称凸関数** ϕ から誘導される: $\phi: \mathbb{R}^n \rightarrow (-\infty, +\infty]$, $\phi(Px) = \phi(x)$ (P : 置換)

逆に, 対称凸関数はユニタリ共役不変な凸関数を誘導する [Davis, 1957]

$\phi: L\text{-平滑} \iff Q: L\text{-平滑}$

$$\text{minimize}_{x \in \text{PD}(n)} \quad Q(df(x)) = \Phi(\tau_{I \rightarrow x}^* df(x)) = \phi(\text{spec}(\tau_{I \rightarrow x}^* df(x)))$$

- $\phi: \mathbb{R}^n \rightarrow (-\infty, +\infty]$, 下半連続, 凸, $\text{int dom } \phi$ で可微分, **対称**: $\phi(Px) = \phi(x)$
- $\Phi: \text{Herm}(n) \rightarrow (-\infty, +\infty]$, $\Phi(X) := \phi(\text{spec } X)$
- $Q: T^* \text{PD}(n) \rightarrow (-\infty, +\infty]$, $Q_x := \Phi \circ \tau_{I \rightarrow x}^*$

最も右の表示は, 形式的には「Euclid 空間上の最適化」

制約集合 $\{\text{spec}(\tau_{I \rightarrow x}^* df(x)) \mid x \in \text{PD}(n)\}$ は陰的に定義され, 一般には凸性も不明.

f が **Kempf–Ness 関数** と呼ばれる種類の関数なら, 上の制約集合は, 閉包を取ると凸多面体になる [Kirwan, 1984, Ness and Mumford, 1984].

(**モーメント多面体 (moment polytope)** と呼ばれる)

→ モーメント多面体上の凸最適化問題は Hadamard 鏡像降下法で解ける!

※対称制約のため, 任意の凸の目的関数が取れるわけではない

$\mathcal{M} := \text{PD}(n_1) \times \cdots \times \text{PD}(n_d)$ 上でも同様の議論が成り立つ：

$$\text{minimize}_{x \in \mathcal{M}} \quad Q(\text{d}f(x)) = \Phi(\tau_{I \rightarrow x}^* \text{d}f(x)) = \phi(\text{spec}(\tau_{I \rightarrow x}^* \text{d}f(x)))$$

※ $\tau_{I \rightarrow x}^* \text{d}f(x) \in \text{Herm}(n_1) \times \cdots \times \text{Herm}(n_d)$, spec は要素ごとのスペクトル

- $\phi : \mathbb{R}^{n_1} \times \cdots \times \mathbb{R}^{n_d} \rightarrow (-\infty, +\infty]$, 下半連続, 凸, $\text{int dom } \phi$ で可微分,
各要素内で対称: $\phi(P_1 x_1, \dots, P_d x_d) = \phi(x_1, \dots, x_d)$ (P_ℓ : 置換行列 ($\ell \in [d]$))
- $\Phi : \text{Herm}(n_1) \times \cdots \times \text{Herm}(n_d) \rightarrow (-\infty, +\infty]$, $\Phi(X) := \phi(\text{spec } X)$
- $Q : T^* \mathcal{M} \rightarrow (-\infty, +\infty]$, $Q_x := \Phi \circ \tau_{I \rightarrow x}^*$

以下では f として, テンソル作用の Kempf–Ness 関数を考える

$\implies \{\text{spec}(\tau_{I \rightarrow x}^* \text{d}f(x)) \mid x \in \mathcal{M}\}$ はエンタングルメント多面体 (entanglement polytope) [Walter et al., 2013] と呼ばれる.

d -テンソル： d 次元配列

($d = 1 \implies$ ベクトル, $d = 2 \implies$ 行列)

$\mathbb{C}^{n_1} \otimes \cdots \otimes \mathbb{C}^{n_d}$: サイズ $n_1 \times \cdots \times n_d$ の複素テンソル全体の空間

自然な内積 $(\cdot, \cdot)$ が入る.

線形写像が自然に作用する：テンソル作用

テンソル $T \in \mathbb{C}^{n_1} \otimes \cdots \otimes \mathbb{C}^{n_d}$ と $g_1 \in \mathrm{GL}(n_1), \dots, g_d \in \mathrm{GL}(n_d)$ に対し,

$$(g_1, \dots, g_d) \cdot T_{i_1, \dots, i_d} := \sum_{j_1, \dots, j_d} (g_1)_{i_1, j_1} \cdots (g_d)_{i_d, j_d} T_{j_1, \dots, j_d}.$$

テンソル $0 \neq T \in \mathbb{C}^{n_1} \otimes \dots \otimes \mathbb{C}^{n_d}$

$1 \leq \ell \leq d$ に対して,

- T の ℓ 番目の **flattening** $T^{(\ell)}$:
 T を $n_\ell \times (n_1 \cdots n_{\ell-1} n_{\ell+1} \cdots n_d)$ 行列とみなしたもの
- ℓ 番目の **one-body marginal** $\mu_\ell(T) := \frac{T^{(\ell)} T^{(\ell)\dagger}}{\|T\|^2} \in \text{PSD}(n_\ell)$

性質: $u_1, \dots, u_d$: ユニタリ $\implies \mu_\ell((u_1, \dots, u_d) \cdot T) = u_\ell \mu_\ell(T) u_\ell^\dagger$

エンタングルメント多面体:

$$\begin{aligned} \Delta(T) &:= \overline{\{(\text{spec}(\mu_1(S)), \dots, \text{spec}(\mu_d(S))) \mid S = (g_1, \dots, g_d) \cdot T, g_\ell \in \text{GL}(n_\ell)\}} \\ &= \overline{\{(\text{spec}(\mu_1(S)), \dots, \text{spec}(\mu_d(S))) \mid S = (x_1, \dots, x_d) \cdot T, x_\ell \in \text{PD}(n_\ell)\}} \end{aligned}$$

非常に複雑な多面体: 一般に頂点や面が指数個存在し, 明示的な表現が困難.

$\Delta(T)$ を計算するアルゴリズム [van den Berg et al., 2025] は, サイズ $4 \times 4 \times 4$ までしか実用的ではない!

テンソル $0 \neq T \in \mathbb{C}^{n_1} \otimes \cdots \otimes \mathbb{C}^{n_d}$ に対し, **Kempf–Ness 関数** は以下で与えられる

$$f_T: \text{PD}(n_1) \times \cdots \times \text{PD}(n_d) \rightarrow \mathbb{R}, \quad f_T(x_1, \dots, x_d) := \log(T, (x_1, \dots, x_d)T)$$

- f_T は測地的凸, d -双対平滑

また実際に微分を計算して平行移動すると,

$$\tau_{I \rightarrow x}^* \text{d}f_T(x_1, \dots, x_d) = (\mu_1(S), \dots, \mu_d(S)) \quad (S := (x_1^{1/2}, \dots, x_d^{1/2}) \cdot T)$$

このスペクトラムを取り閉包を取ると, ちょうど $\Delta(T)$ の表式に一致:

$$\overline{\{\tau_{I \rightarrow x}^* \text{d}f_T(x) \mid x \in \mathcal{M}\}} = \Delta(T)$$

従って, 以下の問題を Hadamard 鏡像降下法で効率的に解くことができる:

$$\min_{p \in \Delta(T)} \phi(p) = \inf_{x \in \mathcal{M}} \phi(\text{spec}(\tau_{I \rightarrow x}^* \text{d}f(x)))$$

ϕ : 下半連続, 凸, $\text{int dom } \phi$ で可微分, $\Delta(T) \subseteq \text{dom } \phi$, 各要素内で対称

$$\min_{p \in \Delta(T)} \phi(p)$$

この最適化問題は，様々なテンソルのランク概念を与える：

- ϕ : ℓ_∞ -ノルム $\implies G$ -stable rank [Derksen, 2022].
- ϕ : 一様分布への ℓ_1 -距離 $\implies$ 非可換ランク [Fortin and Reutenauer, 2004, Garg et al., 2018, Ivanyos et al., 2017, Hamada and Hirai, 2021, Hirai, 2026]
- ϕ : 負のエントロピー $\implies$ Quantum functional [Christandl et al., 2023]

$\theta = (\theta_1, \dots, \theta_d)$: ある確率ベクトル T : d -テンソル

(logarithmic) Quantum functional [Christandl et al., 2023]

$$E_\theta(T) := \max_{(p_1, \dots, p_d) \in \Delta(T)} \sum_{\ell \in [d]} \theta_\ell H(p_\ell) = \sup_{S=(g_1, \dots, g_d) \cdot T} \sum_{\ell \in [d]} \theta_\ell H(\text{spec } \mu_\ell(S))$$

Hadamard 鏡像降下法は $(x_1, \dots, x_d)_k \in \mathcal{M}$ 上のアルゴリズムだが、
対応するテンソル $T_k \propto (x_1^{1/2}, \dots, x_d^{1/2}) \cdot T$ について以下の更新式を得る：

$$T_{k+1} := (\mu_1(T_k)^{-\theta_1/2}, \dots, \mu_d(T_k)^{-\theta_d/2}) \cdot T_k$$

収束レート

$$E_\theta(T) - \sum_{\ell \in [d]} \theta_\ell H(\text{spec } \mu_\ell(T_k)) = \mathcal{O}(k^{-1})$$

これは、Quantum functional を計算する初めての効率的なアルゴリズム

- 鏡像降下法の双対変数による表現を Hadamard 多様体に一般化した
- モーメント/エンタングルメント多面体上の凸最適化に応用可能
ポイント：これらの多面体は非常に複雑だが，我々の手法は多面体を陽に計算する必要が無い

- [Bauschke et al., 2017] Bauschke, H. H., Bolte, J., and Teboulle, M. (2017).
A descent lemma beyond Lipschitz gradient continuity: first-order methods revisited and applications.
[Math. Operet. Res.](#), 42(2):330–348.
- [Beck and Teboulle, 2003] Beck, A. and Teboulle, M. (2003).
Mirror descent and nonlinear projected subgradient methods for convex optimization.
[Operations Research Letters](#), 31(3):167–175.
- [Christandl et al., 2023] Christandl, M., Vrana, P., and Zuiddam, J. (2023).
Universal points in the asymptotic spectrum of tensors.
[Journal of the American Mathematical Society](#), 36(1):31–79.
- [Davis, 1957] Davis, C. (1957).
All convex invariant functions of hermitian matrices.
[Archiv der Mathematik](#), 8(4):276–278.
- [Derksen, 2022] Derksen, H. (2022).
The G-stable rank for tensors and the cap set problem.
[Algebra & Number Theory](#), 16(5):1071–1097.
- [Doğan et al., 2026] Doğan, M. L., Sakabe, K., and Walter, M. (2026).
Convex optimization on moment polytopes: Hadamard mirror descent and efficient algorithms for quantum functionals and other tensor parameters.
- [Fortin and Reutenauer, 2004] Fortin, M. and Reutenauer, C. (2004).
Commutative/noncommutative rank of linear matrices and subspaces of matrices of low rank.
[Séminaire Lotharingien de Combinatoire](#), 52:B52f.
- [Garg et al., 2018] Garg, A., Gurvits, L., Oliveira, R., and Wigderson, A. (2018).
Algorithmic and optimization aspects of Brascamp–Lieb inequalities, via operator scaling.
[Geom. Funct. Anal.](#), 28:100–145.

- [Hamada and Hirai, 2021] Hamada, M. and Hirai, H. (2021).
Computing the nc-rank via discrete convex optimization on $\text{CAT}(0)$ spaces.
[SIAM Journal on Applied Algebra and Geometry](#), 5(3):455–478.
- [Hirai, 2025] Hirai, H. (2025).
Generalized gradient flows in Hadamard manifolds and convex optimization on entanglement polytopes.
[arXiv:2511.12064](#).
- [Hirai, 2026] Hirai, H. (2026).
A scaling characterization of nc-rank via unbounded gradient flow.
[Linear Algebra and its Applications](#), 730:525–545.
- [Hirai and Sakabe, 2024] Hirai, H. and Sakabe, K. (2024).
Gradient descent for unbounded convex functions on Hadamard manifolds and its applications to scaling problems.
In [2024 IEEE 65th Annual Symposium on Foundations of Computer Science \(FOCS 2024\)](#), pages 2387–2402.
- [Ivanyos et al., 2017] Ivanyos, G., Qiao, Y., and Subrahmanyam, K. V. (2017).
Non-commutative Edmonds' problem and matrix semi-invariants.
[Comput. Complex.](#), 26(3):717–763.
- [Jiang et al., 2026] Jiang, J., Shi, L., and Tan, J. (2026).
Mirror descent on Riemannian manifolds.
[Anal. Appl.](#), (S0219530526500521):1–23.
- [Kirwan, 1984] Kirwan, F. (1984).
Convexity properties of the moment mapping. III.
[Inventiones mathematicae](#), 77(3):547–552.
- [Lu et al., 2018] Lu, H., Freund, R. M., and Nesterov, Y. (2018).
Relatively smooth convex optimization by first-order methods, and applications.
[SIAM J. Optim.](#), 28(1):333–354.

- [Nemirovsky and Yudin, 1983] Nemirovsky, A. S. and Yudin, D. B. (1983).
Problem Complexity and Method Efficiency in Optimization.
Wiley, New York, NY.
- [Ness and Mumford, 1984] Ness, L. and Mumford, D. (1984).
A stratification of the null cone via the moment map.
American Journal of Mathematics, 106(6):1281–1329.
- [Rockafellar, 1970] Rockafellar, R. T. (1970).
Convex Analysis.
Princeton University Press, Princeton, NJ.
- [Sakai, 1996] Sakai, T. (1996).
Riemannian Geometry.
American Mathematical Society, Providence, RI.
- [van den Berg et al., 2025] van den Berg, M., Christandl, M., Lysikov, V., Nieuwboer, H., Walter, M., and Zuiddam, J. (2025).
Computing moment polytopes of tensors, with applications in algebraic complexity and quantum information.
In Proceedings of the 57th Annual ACM Symposium on Theory of Computing, STOC '25, pages 756–765. ACM.
- [Walter et al., 2013] Walter, M., Doran, B., Gross, D., and Christandl, M. (2013).
Entanglement polytopes: Multiparticle entanglement from single-particle information.
Science, 340(6137):1205–1208.