YSM-SPIP2011Kansai

機械学習に隠されたJarzynski等式

京都大学大学院情報学研究科

大関　真之

共同研究者：田中　利幸（京大院情報）

Jarzynski等式

Jarzynski等式の意義

- 平衡状態と非平衡過程の橋渡し
  - 平衡状態スタートの動的過程で成立
    - ハミルトン力学、シュレーディンガー方程式（決定論）
    - マスター方程式、ランジュバン方程式（確率過程）

$$\left\langle \exp\left(-\beta W\right)\right\rangle = \exp\left(-\beta \Delta F\right)$$

【非平衡過程に関する平均】　【２つの平衡状態の熱力学量】

Jarzynski等式の意義

- 微視的状態と巨視的世界の橋渡し
  - 左辺の平均は微視的な状態に関する和
    - 全ての実現状態に関して和を取る必要がある
  - 右辺の量は熱力学量
    - 始点と終点で決まる状態量

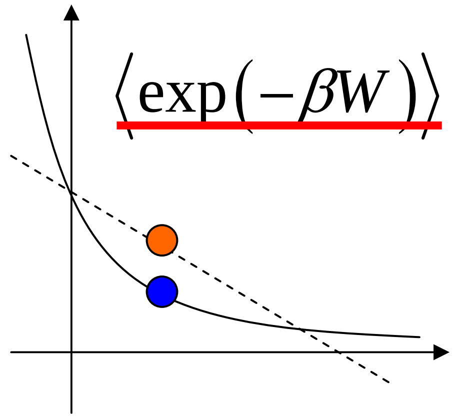
$$\left\langle \exp\left(-\beta W\right)\right\rangle = \exp\left(-\beta \Delta F\right)$$

【非平衡過程に関する平均】　【２つの平衡状態の熱力学量】

Jarzynski等式の意義

- 熱力学第二法則の再現
  - Jensen不等式による評価

$$\left\langle \exp(-\beta W) \right\rangle \ge \exp(\langle -\beta W \rangle)$$



$$\langle W \rangle \ge \Delta F$$

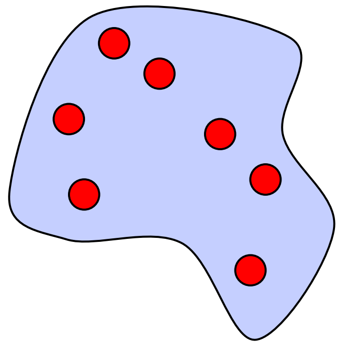
  - Dissipative workが必ず正＝熱力学第二法則

$$\langle W_d \rangle = \langle W \rangle - \Delta F \ge 0$$

Jarzynski等式の成立

- 静止系！のJarzynski等式
  - 位相空間上の点を平衡分布に従い生成させる

$$P(z) = \frac{1}{Z_0}\exp\left(-\beta H_0(z)\right)$$



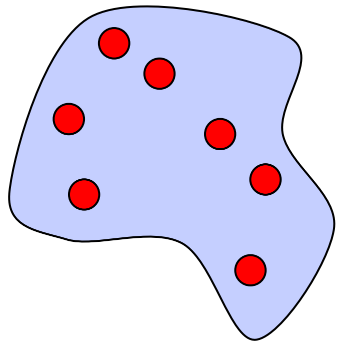
  - 仕事の定義

$$dW = \frac{\partial H}{\partial t}dt = H_{t+1}(z) - H_t(z)$$

Jarzynski等式の成立

- 各点に仕事を乗せる
  - 位相空間上の点それぞれに仕事を割り当てる

$$W=\int_0^T \frac{\partial H}{\partial t}dt = H_T(z)-H_0(z)$$

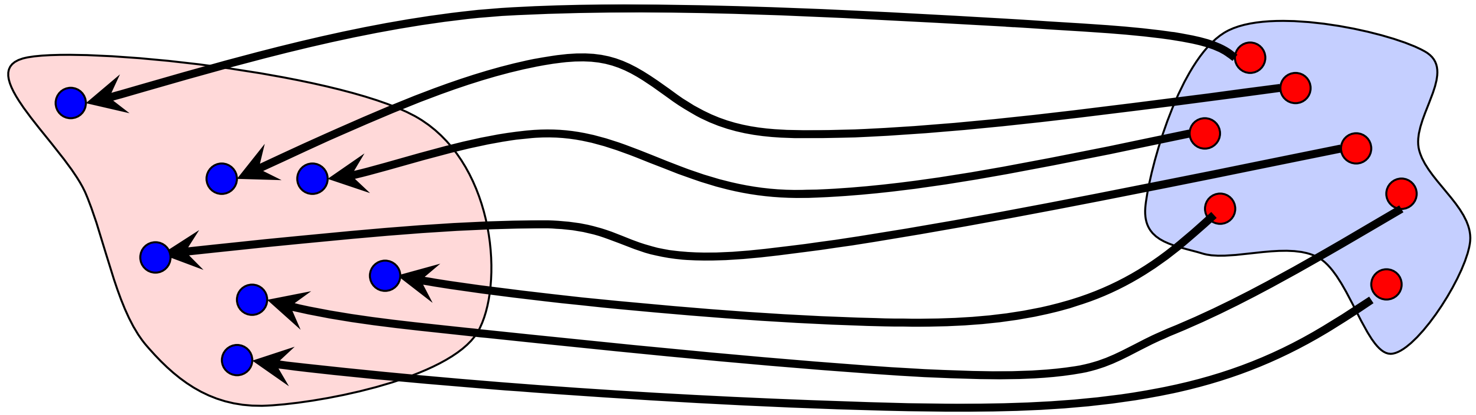


  - Jarzynski等式の左辺＝仕事の重みつき平均

$$\int dz \exp(-\beta W)\frac{\exp(-\beta H_0(z))}{Z_0}=\frac{Z_T}{Z_0}$$

Jarzynski等式の成立

- 孤立系のJarzynski等式

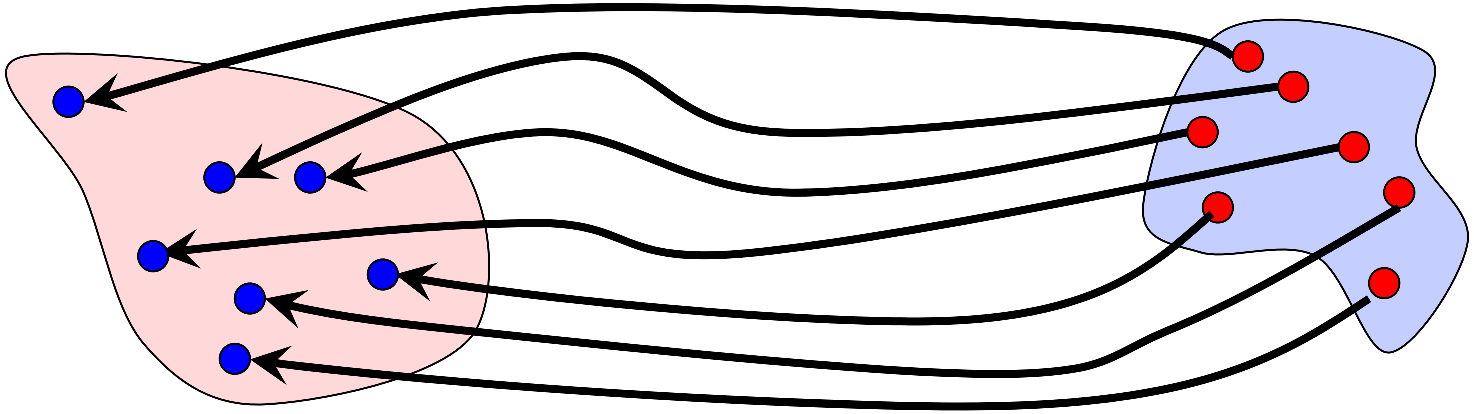


$$P(z) = \frac{1}{Z_0}\exp\left(-\beta H_0(z)\right)$$

- 位相空間上の点からパスが出る（ハミルトン力学）

Jarzynski等式の成立

- 各パスに仕事を乗せる
  - 位相空間上のパスそれぞれに仕事を割り当てる



  - Jarzynski等式の左辺＝仕事の重みつき平均

$$\int Dz \exp(-\beta W)\rho_T$$

Jarzynski等式の成立

- リウビルの定理を使う
  - 終点での平均（パスの平均）は始点での平均

$$\int dz \rho_t \exp(-\beta W) = \int dz \frac{\exp(-\beta H_0(z))}{Z_0} \exp(-\beta W)$$

$$= \int dz \frac{\exp(-\beta H_T(z))}{Z_0} = \frac{Z_T}{Z_0}$$

  - 最終時刻で非平衡分布でもJarzynski等式！

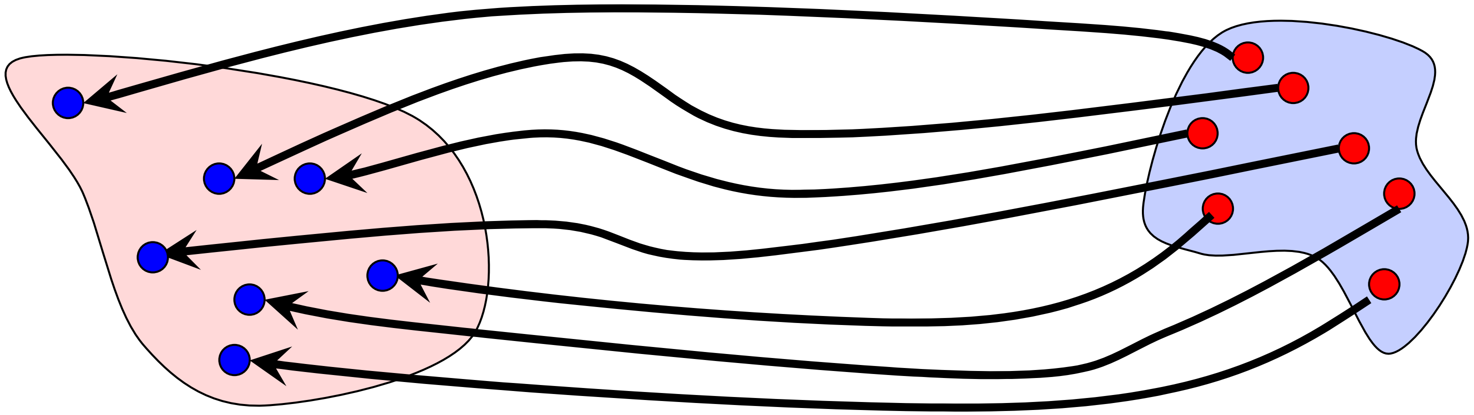
$$\int dz \exp(-\beta W) \rho_T = \frac{Z_T}{Z_0}$$

Jarzynski等式の成立

- リウビルの定理の意味
  - 分布密度の不変性

$$dz_t \rho_t = dz_{t'} \rho_{t'}$$

  - パスの分離可能性，逆に戻せる＝静止系と等価



Jarzynski等式部のまとめ

- 非平衡統計力学：Jarzynski等式
  - 熱力学第二法則の拡張理論
  - 成立は重みつき平均に過ぎない
    - ハミルトン力学など決定論や確率過程でも成立
- 最近の進展
  - 平衡状態からの"距離"とDissipative workの関係 [S. Vaikuntanathan 2009]

$$\langle W_d \rangle \geq D[P(T) \mid P_{eq}(T)] / \beta$$

  - フィードバック制御された仕事 [Sagawa 2010]
  - スピングラス [Ohzeki 2010,11]

機械学習

分類問題

- 自動化された分類システムの開発
  - 電子メールのスパムフィルタ
  - 動画，静止画における顔検出技術
- 抽象化された2値問題で考える
  - 事前に与えられた学習データが存在する
    - 特徴データ $x_i \in X = \{-1,+1\}^{N_D}$
    - 正解を与える2値 $y_i \in Y = \{-1,+1\}$
  - 正解ラベルを正しく推定する優秀な判別器を生成したい

$$y_i f\left(x_i\right) = 1$$

Boosting技術

- 簡単な判別器の組み合わせにより柔軟に対応したい
  - アルゴリズムとしての汎用性の要求
  - 逐次的な改善が行える
  - 過学習の問題は残る

    簡単な判別器＝半分以上は正しく推定できる弱判別器

$$\varepsilon(h;1) < 1/2$$

$$\varepsilon(h;D) = \sum_{i=1}^{N_D} D_i I[y_i \neq h(x_i)] / \sum_{i=1}^{N_D} D_i$$

AdaBoostの特徴

【最終仮説】

$$f_T(x_i)=\mathrm{sign}\left(-\sum_{t=0}^{T}\alpha_t h_t(x_i)\right)$$

- 重みにより難題をうまく判別できるよう更新
  - 適応度と問題の関係による

【正解】

$$\exp(-\alpha_t)$$

【不正解】

$$\exp(\alpha_t)$$

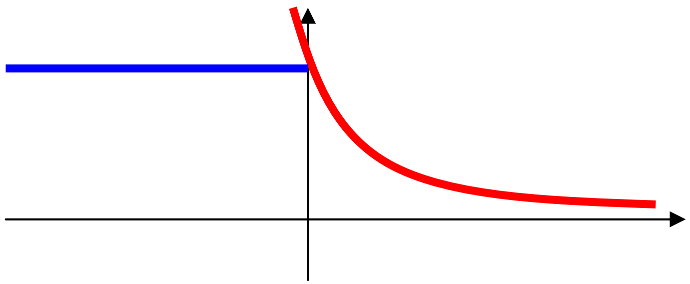
- 最終仮説の誤り確率の単調減少性の保障
  - 弱判別器の性質による

$$\varepsilon(f_T;1)<\prod_{t=1}^{T}\sqrt{2\varepsilon_t(1-\varepsilon_t)}$$

AdaBoostとJarzynski等式

- 最終仮説の誤り確率の評価
  - 弱判別器の性質による

$$\exp(-f) \geq I[f<0]$$



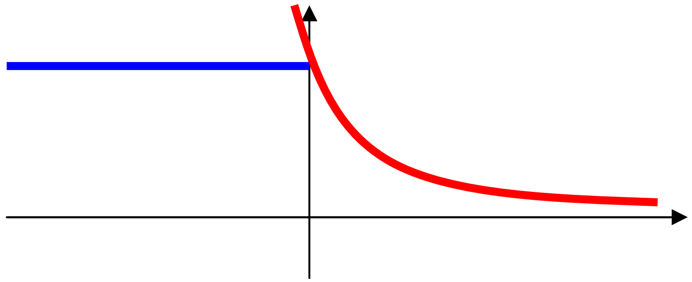
$$\varepsilon(f_T;1)=\sum_{i=1}^{N_D} I[y_i \neq f_T(x_i)]/N_D$$

$$<\sum_{i=1}^{N_D}\exp\left(-y_i\sum_{t=1}^{T}\alpha_t h_t(x_i)\right)/N_D=\prod_{t=1}^{T}\sqrt{2\varepsilon_t(1-\varepsilon_t)}$$

AdaBoostとJarzynski等式

- 最終仮説の誤り確率の評価
  - 弱判別器の性質による

$$\exp(-f) \geq I[f<0]$$



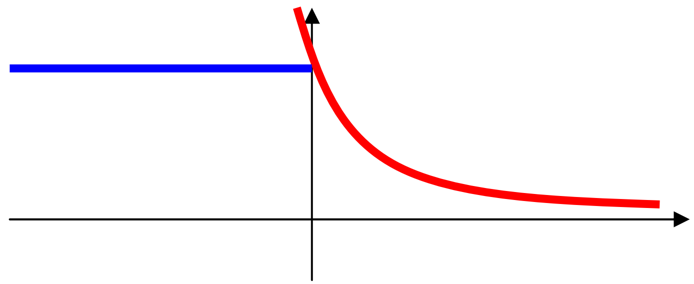
$$\varepsilon(f_T;1) = \sum_{i=1}^{N_D} I[y_i \neq f_T(x_i)] / N_D$$

$$< \sum_{i=1}^{N_D} \exp(-y_i \sum_{t=1}^{T} \alpha_t h_t(x_i)) / N_D = \prod_{t=1}^{T} \sqrt{2\varepsilon_t(1-\varepsilon_t)}$$

  - ん？これはGibbs-Boltzmann分布ではないのか？

AdaBoostとJarzynski等式

$$\exp(-f) \geq I[f<0]$$



- 最終仮説の誤り確率の評価
  - 弱判別器の性質による

$$\varepsilon(f_T;1)=\sum_{i=1}^{N_D} I[y_i \neq f_T(x_i)]/N_D$$

$$<\sum_{i=1}^{N_D}\exp\left(-y_i\sum_{t=1}^{T}\alpha_t h_t(x_i)\right)/N_D=\prod_{t=1}^{T}\sqrt{2\varepsilon_t(1-\varepsilon_t)}$$

  - ん？これはGibbs-Boltzmann分布ではないのか？

$$-\beta H_t(x)=-y_i\sum_{j=1}^{t}\alpha_j h_j(x) \qquad -\beta W_t(x)=-y_i\alpha_t h_t(x)$$

と解釈してみよう

AdaBoostとJarzynski等式

- Jarzynski等式から分かること

$$\left\langle \exp\left(-\beta W\right)\right\rangle = \exp\left(-\beta \Delta F\right); -\beta W_t\left(x\right) = -y_i \alpha_t h_t\left(x\right)$$

  - Jensen不等式（熱力学第二法則）

$$y_i \left\langle \alpha_t h_t\left(x\right)\right\rangle \leq \sum_{t=1}^{T} \log \sqrt{2\varepsilon_t (1-\varepsilon_t)}$$

  - 指数関数と符号関数の不等式

$$\left\langle I[y_i \sum_{t=1}^{T} \alpha_t h_t\left(x\right) < 0] \right\rangle \leq \prod_{t=1}^{T} \sqrt{2\varepsilon_t (1-\varepsilon_t)}$$

AdaBoostとJarzynski等式

- Jarzynski等式から分かること
  - 物理学としては？
    - 実は仕事の過程は新しいフィードバック制御で決まる
      - 各時刻の誤り確率（熱平均）により適応的に次の仕事を決定
      - 自由エネルギーの逐次最小化
  - 学習理論としては？
    - 静止型Jarzynski等式が現行のAdaBoost（発展の脈？）
    - データが変化しても構わない（ハミルトン力学？）
    - 確率過程でデータを生成する場合もAdaBoostは有効（オンライン学習）

$$\left\langle I\left[y_i \sum_{t=1}^{T} \alpha_t h_t(x) < 0\right] \right\rangle \leq \prod_{t=1}^{T} \sqrt{2\varepsilon_t (1-\varepsilon_t)}$$

妄言

- AdaBoostと（静的）Jarzynski等式
  - 期待されること
    - オンライン学習や動的データの判別問題への応用
    - 適応的な仕事によるJarzynski等式

妄言

- AdaBoostと（静的）Jarzynski等式
  - 期待されること
    - オンライン学習や動的データの判別問題への応用
    - 適応的な仕事によるJarzynski等式
- AdaBoostと情報幾何学（早大：村田昇）
  - AdaBoostの幾何学的な観点からの解釈法

妄言

- AdaBoostと（静的）Jarzynski等式
  - 期待されること
    - オンライン学習や動的データの判別問題への応用
    - 適応的な仕事によるJarzynski等式
- AdaBoostと情報幾何学（早大：村田昇）
  - AdaBoostの幾何学的な観点からの解釈法
- Jarzynski等式と進化の統計力学（東大：坂田綾香）
  - 各時刻の熱平均値によって次の変化が決まる意味で同じ

妄言

- AdaBoostと（静的）Jarzynski等式
  - 期待されること
    - オンライン学習や動的データの判別問題への応用
    - 適応的な仕事によるJarzynski等式
- AdaBoostと情報幾何学（早大：村田昇）
  - AdaBoostの幾何学的な観点からの解釈法
- Jarzynski等式と進化の統計力学（東大：坂田綾香）
  - 各時刻の熱平均値によって次の変化が決まる意味で同じ
- Jarzynski等式と情報幾何学…
  - 情報幾何に非平衡分布の計量としての可能性がある

そういえば...

- 平衡状態からの”距離”とDissipative workの関係 [S. Vaikuntanathan 2009]

$$\langle W_d \rangle \geq D[P(T) \mid P_{eq}(T)] / \beta$$

…カルバック＝ライブラー発散が

鍵を握ってましたね？